DOCUMENT RESUME

ED 083 325 UD 013 718

AUTHOR Barabas, Jean, Comp.
TITLE The Assessment of Minority Groups: An Annotated Bibliography. ERIC-IRCD Urban Disadvantaged Series, Number 34, August 1973.
INSTITUTION Columbia Univ., New York, N.Y. ERIC Clearinghouse on the Urban Disadvantaged.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Aug 73
CONTRACT OEC-0-9-420088-2327
NOTE 85p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Achievement Tests; *Annotated Bibliographies; Attitude Tests; Culture Free Tests; *Educational Diagnosis; Intelligence Tests; *Minority Groups; Personality Tests; Predictive Ability (Testing); Racial Differences; Social Differences; *Testing; *Testing Problems; Test Reliability; Test Validity

ABSTRACT

The materials cited here represent information on such diverse but interrelated areas as: methods of assessing achievement, intelligence, personality factors, and attitudes; effects of testing on self concept and employment opportunities; prediction of academic success; reliability and validity of specific tests; criticism of the methods and use of assessment; test construction; use of assessment for educational placement and diagnosis; culture free and culture fair tests; performance differences on tests between majority and minority groups. The educational literature from the early 1960's to the present from the following sources was searched: the ERIC system, the Library of Teachers College, Columbia University, the in-house document collection of the ERIC Information Retrieval Center on the Disadvantaged, and several existing bibliographies. A few pieced dated prior to 1960 were included. The bibliography is arranged alphabetically by author and is followed by a subject index. Each citation is cross-referenced under two or more of the index terms which, hopefully, best describe the material. The bibliography is completely annotated. For those documents which deal only in part or peripherally with minority groups, relevant quotations are given or the relevant sections are described. A fuller annotation is included for those materials which deal totally with minority group assessment. (Author/JM)

FILMED FROM BEST AVAILABLE COPY

ERIC/IRCD URBAN DISADVANTAGED SERIES

Number 34, August 1973

ED 083325

THE ASSESSMENT OF MINORITY GROUPS
AN ANNOTATED BIBLIOGRAPHY

Jean Barabas
Research Assistant

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ERIC INFORMATION RETRIEVAL CENTER ON THE DISADVANTAGED

Horace Mann-Lincoln Institute · Teachers College · Columbia University
New York, New York 10027

UD 013 718

This bibliography is one in the ERIC/IRCD Urban Disadvantaged Series produced by the ERIC Information Retrieval Center on the Disadvantaged under Contract OEC-0-9-420088-2327 between the National Institute of Education (DHEW) and Teachers College, Columbia University. It has been assigned the ERIC/IRCD accession number UD 013 718, and is available without cost from ERIC/IRCD during the period immediately following publication; thereafter, it can be obtained from:

ERIC Document Reproduction Service (EDRS)
Leasco Information Products, Inc. (LIPCO)
P. O. Drawer 0
Bethesda, Maryland 20014

In ordering from EDRS, use the document's assigned ED number. The bibliography will be abstracted in a monthly issue and announced in the semi-annual and annual indexes of Research in Education (RIE). Its ED ordering number will be listed in these indexes or can be obtained by writing to ERIC/IRCD.

This bibliography was prepared pursuant to a contract with the National Institute of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official National Institute of Education position or policy.

August 1973

Table of Contents

Preface	v
Availability Statement	vi
The Bibliography	1
Subject Index	75

Preface

As the compilation of this bibliography proceeded, the complexity of the issue of minority group assessment became apparent. The materials cited here represent information on such diverse and, at the same time, intertwining areas as: methods of assessing achievement, intelligence, personality factors, attitudes; effects of testing on self concept and employment opportunities; prediction of academic success; reliability and validity of specific tests; criticism of the methods and use of assessment; test construction; use of assessment for educational placement and diagnosis; culture free and culture fair tests; performance differences on tests between majority and minority groups.

The educational literature from the early 1960's to the present from the following sources was searched: the ERIC system (both Research in Education and Current Index to Journals in Education), the Library of Teachers College, Columbia University, the in-house document collection of the ERIC Information Retrieval Center on the Disadvantaged, and several existing bibliographies, notably "Research on Assessment for Guidance of Minority Adolescents, An Annotated Bibliography" by Linda M. Staples, Educational Testing Service, Princeton, New Jersey. A few pieces dated prior to 1960 were included.

Because of the overlapping nature of the contents of so many of the cited documents, it was decided not to try to classify them into subgroups. Rather, the bibliography is arranged alphabetically by author and is followed by a subject index. Each citation is cross-referenced under two or more of the index terms which, hopefully, best describe the material. The bibliography is completely annotated. For those documents which deal only in part or peripheraliy with minority groups, relevant quotations are given or the relevant sections are described. A fuller annotation is included for those materials which deal totally with minority group assessment.

There are a number of citations in the bibliography which do not deal directly with minorities. Several general works which touch on theoretical and technical issues which may be related to the assessment of minority groups have been included. For examples see citations 3, 7, and 231. Also included are citations which generally deal with criterion referenced tests. (See the subject index.) A criterion referenced test is one that compares an individual's performance to specified performance standards, as opposed to comparing it with the performance of a norm group. The recent interest, especially in relation to minority groups, in areas such as accountability and performance contracting, where the concern is with how a student does in relation to an outside set of standards, prompted the inclusion of these citations.

Conflicting views and findings are presented in various cited documents. All were included in order to present as complete and objective a bibliography as possible. The views expressed in the cited documents do not necessarily represent the views or policy of the ERIC Information Retrieval Center on the Disadvantaged.

Jean Barabas
Research Assistant

Availability Statement

Unless otherwise noted, documents cited in this bibliography which are accompanied by "ED" numbers are available from:

ERIC Document Reproduction Service (EDRS)
Leasco Information Products, Inc.
P. O. Drawer 0
Bethesda, Maryland 20014

When ordering from EDRS, orders must be in writing stating the ED number; orders under $10.00 must be accompanied by payment. Both microfiche (MF) and hard copy (HC) prices are quoted following the citations. Microfiche is a 4 x 6 inch sheet of film; hard copy is a paper photo copy or facsimile of the original document.

Journal articles and published books cited are not available from the ERIC system. They can be found in many libraries or can be obtained through a bookstore or directly from the publisher.

1. Abramson, Theodore. The Influence of Examiner Race on First-Grade and Kindergarten Subjects' Peabody Picture Vocabulary Test Scores. Journal of Educational Measurement, 6(4): 241-246, Winter 1969.

The tests administered by two white and two Negro examiners to 88 and 113 white and Negro first grade and kindergarten subjects, respectively, in an integrated urban school found a small but statistically significant interaction of the examiner's race and subject's race on test performance for the first-grade but not for the kindergarten.

2. Adler, M. Intelligence Testing of the Culturally Disadvantaged. Some Pitfalls. Journal of Negro Education, 37(4): 364-369, Fall 1968.

"A gifted child from this [minority groups, lower socioeconomic class, poverty] segment of our society may often be lost because of the discrimination inherent in many of our methods of identification. If we are to identify and develop these valuable human resources, greater caution will be necessary in the use of tests as sole measures of academic potential." (p. 369)

3. Aiken, Lewis R. Psychological and Educational Testing. Boston, Mass., Allyn and Bacon, Inc., 1971.

This is a general reference book on psychological and educational testing. Chapter 6, "General Intelligence: Research and Theories," discusses distribution of IQ, the stability of intelligence, social class differences in IQ, and the heredity-environment question. The final chapter discusses current problems and issues in testing such as testing and fair employment and race and differential prediction. A glossary of testing terms, a list of test publishers and a list of references are included.

4. Altus, G. T. Some Correlates of the Davis-Eells Tests. Journal of Consulting Psychology, 20(3): 227-232, 1956.

The Davis-Eeells Games (designed to eliminate some of the "cultural bias" in other intelligence tests) were found to have low correlations with the California Test of Mental Maturity and the California Tests of Reading and Arithmetic. The author concluded from this that "the Games should not be used in place of standard group intelligence tests for predicting school success, grouping pupils for instructional purposes, and similar traditional uses."

5. Anastasi, Anne. Culture-fair Testing. Educational Horizons, 43(1): 26-30, 1964.

A discussion of the difference between "culture-free" and "culture-fair" tests is presented. It is asserted that "culture-free" tests do not exist. The article then proceeds to discuss the rationale for the construction and variety of, and the use of "culture-fair" tests.

6. Anastasi, Anne and de Jesus, Cruz. Language Development and Nonverbal IQ of Puerto Rican Preschool Children in New York City. Journal of Abnormal and Social Psychology, 48(3): 357-366, 1953.

It is asserted in this article that because of being unfamiliar with English, Puerto Rican children produce a sort of "psychological insulation" to whatever goes on in school. Passivity and unresponsiveness become the child's characteristic response to school and performance on psychological tests is seriously handicapped by this passivity. After prolonged exposure to school environment, the deficiency on test performance increases.

7. Anastasi, Anne. Psychological Testing. 2nd edition. New York, N. Y., The MacMillan Company, 1968.

This book is a basic text designed to teach the reader to evaluate psychological tests and interpret test results. The use of tests, test reliability, test validity, specific tests for ability, aptitude, achievement, and the measurement of personality traits are discussed.

8. Anastasi, Anne. Psychological Tests: Uses and Abuses. Teachers College Record, 62(5): 389-393, February 1961.

This article discusses the use of tests in intergroup comparisons. When comparing different groups, the only question tests can directly answer is: "How do these groups differ under existing cultural conditions?" To improve the use of tests in intergroup comparisons, "first, describe differences as they exist in the present; second,...investigate their origins in past events; third,...predict future outcomes." (p. 393)

9. Anastasi, Anne and Cordova, F. A. Some Effects of Bilingualism Upon the Intelligence Test Performance of Puerto Rican Children in New York City. Journal of Educational Psychology, 44(1): 1-19, January 1953.

The results of a study involving the testing of Puerto Rican children in both Spanish and English with the Cattell "Culture Free" Test are discussed The overall performance of the group fell considerably below the test norms reported by Cattell. Reasons for this are the very low socioeconomic level of the Puerto Rican children, their bilingualism which makes them deficient in both languages, their extreme lack of test sophistication, and their poor emotional adjustment to the school situation. "Insofar as this maladjustment itself appears to have arisen from the children's severe language handicap during initial school experiences, a solution of the language problem would seem to be a necessary first step for the effective education of migrant Puerto Rican children." (p. 17)

10. Anastasi, Anne. Standardized Ability Testing. In: Handbook of Research Methods in Child Development edited by Paul H. Mussen. New York, N. Y., John Wiley, 1960. pp. 456-486.

The section entitled "Cultural Factors in Test Scores" (pp. 479ff.) discusses difficulties in administering the interpreting tests for different cultures or subcultures and discusses culture-free tests.

11. Angoff, William H. and Ford, Susan F. Item-Race Interaction on a Test of Scholastic Aptitude. Princeton, N. J., Educational Testing Service, 1971. 26p. ED 058 279 (MF-$0.65; HC-$3.29).

Several samples of black and white students were drawn from the 1970 PSAT administration in Georgia and studied for item x race interaction on both the verbal and mathematical sections of the test. When subsamples of candidates were drawn from their respective racial groups, matched on mathematical for the study of verbal items and matched on verbal for the study of mathematical items, there was an observable decrease in the size of the item x race interaction, suggesting that one factor contributing to that interaction was simply the difference in performance levels on the test shown by the two races. Further analyses demonstrated a moderate item x group interaction for blacks native to different cities and moderate item x group interaction for blacks native to areas of different population density.

12. Arnold, Richard D. Reliability of Test Scores for the Young Bilingual Disadvantaged. Reading Teacher, 22(4): 341-345, January 1969.

The study indicated that the Inter-American English Test of Reading and the Metropolitan Achievement Tests were quite reliable when applied to the particular disadvantaged bilingual children studied [children who had been involved in a research program with an emphasis on extensive oral language development] provided an appropriate difficulty level is administered.

13. Arvey, Richard D. Some Comments on Culture Fair Tests. Personnel Psychology, 25(3): 433-448, Autumn 1972.

Culture fair tests should be examined and reviewed closely before being used and should not be regarded as the answer to testing the culturally disadvantaged.

14. Back, Zellig. Devaluing the IQ. NJEA Review, 45(2): 16-18, October 1971.

Test reliability and test bias for intelligence tests is discussed in relation to cultural differences among the populations tested.

15. Barclay, J. R. Testing for Higher Education, Cultural Perspective and Future Focus. Student Personnel Series No. 6. Washington, D.C., American Personnel and Guidance Association, 1965. ED 017 031 (Available from American Personnel and Guidance Assn., 1605 New Hampshire Ave., N.W., Washington, D.C. 20009, $1.50).

This document is mainly constructed for high school personnel, college admissions officers, counselors and all other interested parties, to shed some light on the use of tests as predictors of college success, the clarification of criteria used to evaluate effective human behavior in this setting and the broader implications of cultural and philosophical dimensions of testing and test usage.

16. Barritt, L. S. Intelligence Test and Educationally Relevant Measurements. Ann Arbor, Mich., University of Michigan, Center for Research in Language and Behavior, 1967.

This report challenges the relevance of intelligence tests on the following

points: 1) tests which merely predict the likelihood of future success do not provide useful information on the performance of the individual and 2) since intelligence is not clearly defined, interpretation of such scores may prove to be misleading.

17. Bernardoni, L. C. A Culture Fair Intelligence Test for the Ugh, No, and Oo-La-La Cultures. Personnel and Guidance Journal, 42: 554-559, 1964.

"The testing of intelligence necessitates testing behavior and behavior devoid of a cultural context is impossible to conceptualize. For a test to be completely fair to a culture, it must consider the values of that culture." (p. 557)

18. Bloom, Benjamin S. et al. Handbook on Formative and Summative Evaluation of Student Learning. New York, N. Y., McGraw-Hill, 1971.

The art and science of evaluating student learning is presented in detail to assist the classroom teacher in its application and the consequent improvement of both the teaching and learning processes. The specialized chapters are intended to help teachers find ways of using evaluation to improve learning in their own specific area. The entire book is related to Bloom's earlier work on the Taxonomy of Educational Objectives. Each chapter is supported by a comprehensive reference list.

19. Bloom, Benjamin S. Stability and Change in Human Characteristics. New York, N. Y., John Wiley and Sons, Inc., 1964.

"The analysis of the relationships between gains and initial and final status raises serious questions about grading practices which are based on status measures rather than gain or growth measures. The discrepancies between grades based on status measures and grades based on gain measures are likely to be rewarded or punished for characteristics they possess at the beginning of a school term, whereas the gains they make during the school term are largely ignored. Thus two individuals may make equal progress during a given school year but be given very different grades at the end of the year because they started at different points at the beginning of the year. The consequences of repeated success or failure over several school years must surely have major effects on the individual's view of himself and his attitudes toward school and school learning." (p. 128)

20. Bonner, Mary W. and Belden, Bernard R. A Comparative Study of the Performance of Negro Seniors of Oklahoma City High Schools on the Wechsler Adult Intelligence Scale and the Peabody Picture Vocabulary Test. Journal of Negro Education, 39(4): 354-358, Fall 1970.

The purpose of this study was to compare the performance of Negroes on 2 instruments which are used to measure intelligence, only one of which (WAIS) included them in its standardization procedure. One of the findings was that the I.Q. estimated by use of the PPVT is somewhat lower than that estimated by the WAIS.

21. Bordie, John G. Language Tests and Linguistically Different Learners: The Sad State of the Art. Elementary English, 47(6): 814-828, October 1970.

This paper discusses four questions: 1) To what extent are currently available measuring techniques useful for identifying the characteristics of linguistically different learners? 2) Are they helpful in planning instructional strategies? 3) How might the learning potentialities of linguistically different learners be measured? 4) What are the high priority test needs? Specific tests are individually discussed.

22. Borup, Jerry H. The Validity of American College Test for Discerning Potential Academic Achievement Levels--Ethnic and Sex Groups. Journal of Educational Research, 65(1): 3-6, September 1971.

This study demonstrates that high school quarter rankings are better indicators of potential college achievement than the American College Test and that the test has a built-in sex and ethnic bias.

23. Boyd, Joseph L., Jr. and Shimberg, Benjamin. Handbook of Performance Testing: A Practical Guide for Test Makers. Princeton, N. J., Educational Testing Service, 1971. 189p. ED 052 220 (MF-$0.65; HC-$6.58).

Particularly concerned with fulfilling the needs of individuals involved in manpower development and training programs, the nature of and need for performance testing is discussed.

24. Briere, Eugene J. and Brown, Richard H. Norming. Tests of ESL Among Amerindian Children. 1971. 14p. ED 052 667 (MF-$0.65; HC-$3.29).

This paper describes the activities to develop norms for the interpretation of tests designed to indicate proficiency in English for Amerindian children attending grades 3 through 6 in the Bureau of Indian Affairs' Schools. The objectives of the test battery are 1) to identify the Amerindian child who needs special training in English and to determine the proper level of intensity of English training; 2) to provide the classroom teacher with specific linguistic information for each child in each language group which could be used as a diagnostic guide for teaching methods or materials; 3) to provide a means of assessing the merit of various English programs.

25. Brown, Roscoe C., Jr. and Henderson, Edward H. A Discriminant Analysis of Variables Used to Select Students from Disadvantaged Backgrounds. 1966. 11p. ED 024 716 (MF-$0.65; HC-$3.29).

The variables of intellectual functioning, personality and social adjustment, and personal goals and motifation were measured by standard tests, school records, a personal data sheet, and a structured interview with a psychologist counselor. The interviewers' recommendations appear to be the major variable differentiating the selected and rejected groups of students.

26. Bruininks, Robert H. and Lucker, William G. Change and Stability in Correlations Between Intelligence and Reading Test Scores Among Disadvantaged Children. Journal of Reading Behavior, 2(4): 295-305, Fall 1970.

"It...appears that IQ scores, even on individually administered tests, are rather poor predictors of early reading performance among disadvantaged children, but reasonably good predictors of later reading performance." (p. 301)

27. Bruininks, Robert H. and Feldman, David H. Creativity, Intelligence, and Achievement Among Disadvantaged Children. Psychology in the Schools, 7(3): 260-264, July 1970.

The major finding of the study was that Torrance's tests of creativity acted as a suppressor variable, increasing the relation between IQ and achievement among the disadvantaged children sample.

28. Bryan, Miriam M. Ability Grouping: Status, Impact, and Alternatives. Princeton, N. J., ERIC Clearinghouse on Tests, Measurement, and Evaluation, 1971. 15p. ED 052 260 (MF-$0.65; HC-$3.29).

Misuses of standardized tests in grouping procedures are fully discussed. It is contended that grouping results in a trend toward improved achievement in superior groups and poorer achievement in average or low groups. In addition, it tends to reinforce favorable self concepts in those assigned to high achievement groups and unfavorable self concepts in those assigned to low achievement groups. There seems to be a negative effect with regard to ethnic and socioeconomic separation.

29. Bucky, Steven F. and Banta, Thomas J. Racial Factors in Test Performance. Developmental Psychology, 6(1): 7-13, January 1972.

The purpose of the study was to determine a) whether there are differences in the way in which Negro and white experimenters interact with Negro and white subjects; and b) whether such differences influence performance on tests of motor impulse control, reflectivity, innovative behavior, and curiosity.

30. Caldwell, Mark B. and Knight, David. The Effect of Negro and White Examiner on Negro Intelligence Test Performance. Journal of Negro Education, 39 (2): 177-179, Spring 1970.

Results of the study indicate that the race of the examiner is not a critical variable in test performance relating to Negro students.

31. Campbell, Joel T. et al. Prediction of Job Performance for Negro and White Medical Technicians. A Regression Analysis of Potential Test Bias: Predicting Job Knowledge Scores from Aptitude Battery. Princeton, N. J., Educational Testing Service; Washington, D. C., Civil Service Commission, 1969. 16p. ED 034 235 (MF-$0.65; HC-$3.29).

All aptitude tests were found to be culturally biased but 7 displayed bias in direction of the whites. On these 7, for the predictor scores, white technicians were likely to score better on the Job Knowledge Test than did Negroes of the same predictor score.

32. Campbell, Joel. Testing of Culturally Different Groups. Research and Development Report 63-4, No. 14. Princeton, N. J., Educational Testing Service, 1964. 22p.

An investigation was undertaken to examine data from several studies concerned with predicting the performance of Negroes and others from deprived backgrounds. The following major conclusions were drawn: cultural deprivation will affect test performance adversely; remedial efforts can improve test performance, although the limits of this improvement has not been established; tests of verbal and arithmetic ability are effective predictors of academic grades in both white and Negro colleges.

33. Carver, Ronald P. An Experiment that Failed: Designing an Aural Aptitude Test for Negroes. College Board Review, 70: 10-14, Winter 1968-1969.

The attempt to design a "culture free" aptitude test for Negro students is described. Reasons for the failure to produce a successful instrument are discussed.

34. Cazden, Courtney B. Subcultural Differences in Child Language: An Interdisciplinary Review. Merrill-Palmer Quarterly of Behavior and Development, 12(3): 185-219, 1966.

"The issue of 'dialect-fair' scales of language development may become as significant in the future as that of 'culture-fair' tests of intelligence has been in the past...ideally, a child's language development should be evaluated in terms of his progress toward the norms for his particular speech community." (p. 204)

35. Chandler, John T. and Plakos, John. Spanish-speaking Pupils Classified as Educable Mentally Retarded. Integrated Education, 7(6): 28-33, November-December 1969.

The results of this investigation indicate that many Mexican-American pupils may have been placed in EMR classes solely on the basis of performance on an English language IQ test when they lack facility and understanding of English. However, when these pupils were tested in Spanish, their performance rose in many cases to above the cut-off level of EMR.

36. Charters, W. W., Jr. Social Class and Intelligence Tests. In: School Children and the Urban Slum, A Book of Readings in Social Psychology for Teachers edited by Joan I. Roberts. New York, N. Y., Hunter College, 1965. pp. 105-114.

This essay discusses the development of culture-free tests.

37. Chase, Clinton I. and Pugh, Richard C. Social Class and Performance on an Intelligence Test. Journal of Educational Measurement, 8(3): 197-202, Fall 1971.

No item-social class interaction appeared when one lower-class group and one middle-class group of children who had achieved comparable total scores on a group intelligence test were compared. A given score on an intelligence test appears to represent the same skills for one social class as it does for another.

38. Chauncey, Henry and Dobbin, John E. Testing: Its Place in Education Today. New York, Harper and Row, 1963.

This book is an overview of current educational testing developments in schools and colleges. It provides a brief history of testing and, among others, chapters on achievement tests, test selection, and tests used in guidance work.

39. Chovan, W. L. and Hathaway, M. L. The Performance of Two Divergent Groups of Children on a Culture-Free Test. Journal of School Psychology, 8(1): 66, 1970.

This brief report discusses the investigation of the nature and extent of cultural differences that operate in the measures of intelligence recall and motor coordination between randomly selected, divergent socioeconomic groups of five and one-half-year-old children.

40. Cicirelli, Victor G. A Note on the Factor Analysis of Disadvantaged Children's Illinois Test of Psycholinguistic Abilities and Achievement Test Scores. Journal of Experimental Education, 41(1): 5-8, Fall 1972.

Two main factors were identified: an achievement factor and a language factor.

41. Cleary, T. Anne. Test Bias: Prediction of Grades of Negro and White Students in Integrated Colleges. Journal of Educational Measurement, 5: 115-124, 1968.

The results of this study show that in the three colleges studied, there was little evidence that the Scholastic Aptitude Test is biased as a predictor of college grades, and when high school grades or rank-in-class are used in addition to the SAT as predictors, the degree of positive bias for the Negro students increases.

42. Cleary, T. Anne. Test Bias--Validity of the Scholastic Aptitude Test for Negro and White Students in Integrated Colleges. Princeton, N. J., Educational Testing Service, 1966. 26p. ED 018 200 (MF-$0.65; HC-$3.29).

For this research, a test was said to be biased for members of a subgroup of the population if, in the prediction of a criterion for which test was designed, consistent nonzero errors of prediction are made for members of the subgroup. Samples of Negro and white students from three integrated colleges were studied. In the two Eastern colleges, no significant differences in the regression lines were found. In one college in the Southwest, significant differences were found, but it was the Negro scores that were overpredicted. Thus, in one of the three schools, the scholastic aptitude test was found to be slightly biased, but in favor of the Negro student.

43. Clarke, Eunice A. Exploratory Analysis of an Assessment Procedure for the Placement and Regrouping of Children from Depressed Areas in and within the Ungraded Primary Unit. Philadelphia, Pa., Temple University, 1968. 71p. ED 019 389 (MF-$0.65; HC-$3.29).

"The key factor in deciding placement was found to be the testing program of the Central Administration..." (p. 1)

44. Clark, Kenneth B. and Plotkin, Lawrence. The Negro Student at Integrated Colleges. New York, N. Y., National Scholarship Service Fund for Negro Students, 1963. 59p. ED 011 268 (MF-$0.65; HC-$3.29).

(cont.)

An outstanding finding of a 5-year follow-up study of NSSFNS done on Negro students at interracial colleges reported in this document was the relatively low drop-out rate (about one-half the national average for white and Negro students at segregated schools). Yet the predictive value of precollege test scores was not high, in terms of college grades. The study indicates that motivational factors are probably more important than test scores in the demonstrated superiority of Negro students in completing college.

45. Clark, Phillip I. et al. The Use, Misuse, and Abuse of Tests. New York, N. Y., Harcourt Brace Jovonovitch, Inc.; Providence., R. I., Rhode Island State Agency for Elementary and Secondary Education, 1970. 35p. ED 051 281 (MF-$0.65; HC-$3.29).

The New England Association for Measurement and Evaluation in Guidance (NEAMEG) Conference on Measurement in Education was designed to (1) provide a forum for the examination and discussion of vital issues related to measurement and evaluation; (2) facilitate communication among educators from various disciplines and levels of education within the New England region, and to encourage their active involvement in "attacking" current identified problems and concerns relating to the use of tests and other evaluative devices; and (3) stimulate the development of a series of position papers stating the views of the professional members of the NEAMEG as a group, which may serve as guidelines for education. The proceedings include: "Innovative Test Usage for Individual Pupil Growth," Philip I. Clark; "National Assessment," Thomas R. Knapp; "State Testing Programs," Paul B. Campbell; "Testing the Disadvantaged," Lenore A. DeLucia; "Computerization in Relation to Testing and Evaluation," James R. Baker; "Testing and its Relevance to the Seventies," Thomas Burns; "Federally Funded Programs," Thomas Burns; "Disclosure of Test Results," Thomas P. Nally; "Norms: Fact or Fancy," Walter N. Durost; "Tests: Who or What is Being Evaluated," C. Thomas Skoggs; and "The [illegible]n Report," Paul B. Campbell. A summary of the discussion by the reactors to each presentation follows each paper.

46. Cole, Spurgeon and Hunter, Mildred. Pattern analysis of WISC Scores achieved by Culturally-Disadvantaged Children. Psychological Reports, 29(1): 191-194, August 1971.

An analysis of the WISC profile for culturally disadvantaged children found no evidence to suggest that cultural deprivation has less detrimental effect on performance than on verbal or abstract skills.

47. Cotnam, John Dale. Variance in Self-Report Measures of Disadvantaged Young Adults as a Function of Race and Stated Purpose of Testing. Ph.D. Thesis. Rochester, N. Y., University of Rochester, 1969. 158p. (Available from University Microfilms, 300 N. Zeeb Rd., Ann Arbor, Michigan 48106; Order No. 70-2928).

This study investigated differential effects of students' race, examiners' race, and stated purpose of testing upon self-report measures of disadvantaged young adults. It was hypothesized that examiners of unlike race would represent a "threat" to the subjects. Black and white examiners represented different demand characteristics to subjects of both races, but there was no evidence that examiners of unlike race were a threat to the subjects of either race.

48. Costello, Joan. The Effects of Pretesting and Examiner Characteristics on Test Performance of Young Disadvantaged Children. Chicago, Ill., Institute for Juvenile Research, 1970.

This study reports the effects of 1) pretesting, 2) race of examiner, and 3) familiarity of examiner on the Stanford Binet and Peabody Picture Vocabulary Test scores of 62 preschool black children. Test conditions were carefully controlled for variations not under study. None of these 3 variables had an effect on test performance. These non-significant findings raise questions of significance for decision-making, i.e., is the general atmosphere of testing more powerful than any one variable, or are there unique interactions among variables which may potentiate or neutralize the negative impact of any one examiner variable?

49. Costello, Joan and Ali, Faizienisa. Reliability and Validity of Peabody Picture Vocabulary Test Scores of Disadvantaged Preschool Children. Psychological Reports, 28(3): 755-760, June 1971.

Quite modest validity was suggested by correlations of the standard Peabody test with the other psychological tests assessing intellectual behaviors and with teachers' ratings of several classroom verbal behaviors. While Form A of the Peabody could be used as a first approximation in a continuing assessment program, scores cannot be considered alone for either intellectual or language evaluation.

50. Darlington, Richard B. Another Look at Cultural Fairness. Journal of Educational Measurement, 8(2): 71-82, Summer 1971.

Four definitions of "cultural fairness" are examined and found to be not only mutually contradictory, but all based on the false view that optimum treatment of cultural factors in test construction or test selection can be reduced to completely mechanical procedures.

51. Davis, Allison. Social-Class Influences Upon Learning. Cambridge, Mass., Harvard University Press, 1948. Reprint 1965.

The problems of testing the "mental behavior" of children from various socioeconomic backgrounds are discussed and the construction of tests which would eliminate cultural bias is explored.

52. Davis, W. M., Jr. Are There Solutions to the Problems of Testing Black Americans? Paper presented at the 79th Annual Convention of the American Psychological Association, 1971.

The article discusses the problem of current I.Q. tests that are not measuring inherent intelligence and the fact that a great majority of the tests were not and have not yet been standardized using a black sample. The author gives a list of solutions with emphasis placed on the development and uses of tests adequately standardized on both black and white populations and the development of a more systematic method for adequately measuring socioeconomic levels as it applies particularly to blacks within our culture.

53. Denmark, Florence L. et al. The Effect of Ethnic and Social Class Variables on Semantic Differential Performance. 1969. 12p. ED 036 576 (MF-$0.65; not available in HC).

This report describes a study instituted to determine whether attitude scales standardized on a white population falter when applied to black subjects. The subjects were 656 white and black college students, 457 of these being middle class white students, 29 middle class black students, and 70 lower class black students. Eight semantic differential scales were administered, six of which were related to studying, learning, lectures, group discussion, and college graduates. The remaining two were related to self-perception: real self and ideal self. Findings of the study suggested that the structure of per-

ceptions of concepts related to higher learning differs for blacks and whites, as well as for lower class and middle class blacks. The findings thus question the universality of Osgood's semantic space (Osgood, Suci, and Tannenbaum, 1957) by suggesting that adjectival measures of the same factors vary across ethnic groups.

54. Droege, Robert G. Alternatives to a Moratorium on Testing. 1971. 14p. ED 053 199 (MF-$0.65; HC-$3.29).

The alternative suggested is comprised of three elements: 1) Working toward changes in attitudes of people and institutions to directly attack discriminatory applications of tests; 2) Emphasizing test validation and development by new test instruments oriented to the needs of disadvantaged applicants; and, 3) Eliminating discriminatory selection procedures as stated in Title VII of the Civil Rights Act.

55. Dubin, Jerry A. et al. Speed and Practice: Effects on Negro and White Test Performance. Journal of Applied Psychology, 53(1, pt. 1): 19-23, February 1969.

This research investigated the hypotheses that a) extra pretest practice, b) extra testing time, and c) extra practice and extra testing time would improve the mental ability test performance of Negroes more than whites. Although both races and both socio-economic classes improved performances as the testing procedure became more lenient, all groups profited to a comparable extent. Implications are that the testing procedure itself does not discriminate between racial groups nor between culturally advantaged and disadvantaged subjects.

56. Duggan, J. M. and Hazlett, P. H., Jr. Predicting College Grades. New York, N. Y., College Entrance Examination Board, 1963.

This is a workbook for college officers who want to use a method of combining objective test scores and high school grades to predict an applicant's ability to perform freshman academic work. The workbook is intended to be a non-technical guide that can be used by persons with no statistical training. It presents procedures and computations that enable one to develop a prediction equation that can be used to estimate a student's probable freshman grade average.

57. Dyer, Henry S. Issues in Testing. 1969. 13p. ED 053 164 (MF-$0.65; HC-$3.29).

Certain concepts that are sometimes confused in discussions on testing socially disadvantaged children are clarified and a history of testing beginning with Binet is presented. Finally, five problems in using tests to evaluate educational programs for the disadvantaged are considered.

58. Easterling, Ross E. and Stevens, Francis A. Reading Achievement of Lower Socioeconomic Level Mexican American Students. University Park, N. M., New Mexico State University, ERIC Clearing House on Rural Education and Small Schools, 1972. 19p. ED 068 221 (MF-$0.65; HC-$3.29).

The purpose of the study was to determine if measures of first grade readiness, scholastic aptitude, and reading achievement were significant predictors of reading achievement at the end of third grade for Mexican American students from 2 lower socioeconomic levels. Students (47 boys and 47 girls) who had completed their third year in the Sustained Primary Program for Bilingual Students in Las Cruces, New Mexico, were placed in 9 subgroups. The predictor variables were the 13 tests from which the prediction of third grade reading achievement was made. The data were analyzed by 2 procedures: (1) a multiple regression analysis to determine the relationship between a combination of the 13 tests utilized for prediction and third grade reading achievement and (2) a stepwise multiple regression analysis which identified the individual contribution of each of the 13 tests to the prediction of third grade reading achievement. Findings indicated that a significant multiple correlation existed between the 13 first grade tests used and third grade reading achievement at the .05 level of confidence for (1) total sample of boys, (2) total sample of Social Class IV girls. The report reviewed literature dealing with the effects of social class and/or minority group membership upon test performance and the effect of sex differences on school achievement test scores.

59. Educational Testing Service, Princeton, N. J. Proceedings of the 1970 Invitational Conference on Testing Problems. 1971. 183p. ED 052 264 (MF-$0.65; HC-Available from Educational Testing Service, Princeton, N. J. 08540).

The conference theme was "The Promise and Perils of Educational Information Systems," defined as collections of test data on knowledges, skills, interests, and attitudes maintained for the purpose of educational decision making. Topics covered were: "Longer Education: Thinner, Broader, or Higher"

(Fritz Machlup); "Testing: Americans' Comfortable Panacea" (Theodore R. Sizer); "Social and Cultural Change and the Need for Educational Information: The Futurist's View" (Herman Kahn); "School Testing to Test the Schools" (Richard M. Jaeger); "National Assessment: (Robert E. Stake); "Bayesian Considerations in Educational Information Systems" (Melvin R. Novick); "Temporal Changes in Treatment-Effect Correlations: A Quasi-Experimental Model for Institutional Records and Longitudinal Studies" (Donald T. Campbell); "Higher Education for Whom? At Whose Cost?" (Carl Kaysen); "Social Accounting in Education: Reflections on Supply and Demand" (David K. Cohen); "Ethical and Legal Aspects of the Collection of Educational Information" (David A. Goslin); and "Test Information as a Reinforcer of Negative Attitudes Toward Black Americans" (Elias Blake, Jr.).

60. Educational Testing Service, Berkeley, Calif. and Princeton, N. J. Western Regional Conference on Testing Problems: Proceedings 1970. 1970. 70p. (Available from Educational Testing Service, Berkeley, Calif.; $2.00).

Literature relevant to a basic problem of testing is reviewed in "Test Bias, An Old Problem Revisited." The need for culture-fair tests and even the need for consideration of test bias is questioned.

61. Eells, Kenneth et al. Intelligence and Cultural Differences. Chicago, Ill., The University of Chicago Press, 1951.

The purpose of this study was to identify 1) those kinds of test problems on which children from high socioeconomic backgrounds show the greatest superiority and 2) those kinds of test problems on which children from low socioeconomic backgrounds do relatively well. On the basis of these data, an attempt was made to assess the importance of various factors as possible explanations for the differences in I.Q.'s found for children from different kinds of cultural background. These factors are: genetic ability, developmental factors, cultural bias in test items, test motivation, test work habits or test skills. It was concluded that there was no conclusive proof that any one or any particular combination of these factors was chiefly responsible for status differences. Implications of this study for the construction of new intelligence tests is discussed.

62. Eeells, Kenneth. Some Implications for School-Practice of the Chicago Studies of Cultural Bias in Intelligence Tests. Harvard Educational Review, 23(4): 284-297, 1953.

The meaning of "cultural bias" in intelligence tests and the implications for all the levels of education are discussed. Caution in the use of results of traditional test scores, the need for new intelligence tests, and the development of individual potentialities are considered.

63. Eisenberg, Leon. Child Psychiatry: The Past Quarter Century. Cambridge, Mass., Harvard University, Medical School, 1968. 29p. ED 027 951 (MF-$0.65; HC-$3.29).

Concern for the poor and the black was dormant for too long in the past quarter century; it is in such areas as poverty and racism that serious psychological and organic problems in children occur. Specifically, research in child development must be concerned with many important factors, including the "test bias" in interpreting results of achievement tests.

64. ERIC Clearinghouse on Tests, Measurement, and Evaluation, Princeton, N. J. Test Bias: A Bibliography. 1971. 12p. ED 051 312 (MF-$0.65; HC-$3.29).

This bibliography, which lists selected articles, research reports, monographs, books, and reference works related to test bias, is limited to material which deals directly with the question of test bias--such as research reports of commentaries. References listed cover the period from 1945 to the present, and are grouped under two main headings: General Educational References and Employment References.

65. Farr, James L. et al. Ethnic Group Membership as a Moderator of the Prediction of Job Performance. Personnel Psychology, 24(4): 609-636, Winter 1971.

This article discusses performance on tests of black and white employees on selected civil service and private industry jobs. Test results are then contrasted with actual job performance according to specific criteria and supervisors' evaluations. No tests used permitted valid prediction of job performance for both racial groups.

66. Felsenthal, Norman Allan. Racial Identification as a Variable in Instructional Media. Ph.D. Dissertation. Ames, Iowa, University of Iowa, 1969. 165p.

One aspect of reference group theory, racial identification, was studied, and the effects of this identification on the acquisition of knowledge and the formation of attitudes were examined. Three hypotheses were invetigated: (1) Children who perceive an instructional tape narrator to be a member of their own race will score significantly higher on a retention test than will children who perceive a narrator to be a member of a different race; (2) Children who perceive the narrator to be a member of their own race will rate that narrator higher on the various dimensions of source credibility and will also rate as more acceptable the content the narrator discusses than will subjects who perceive the narrator to be of a different race; (3) The effects anticipated in the first two hypotheses will be greatest among students who are low achievers and least among students who are high achievers. Eighth grade students served as subjects; 40% were Negro; the balance were white. All 256 subjects were exposed to a sound-slide presentation; then subjects were given three successive tests. All criterion measurements were subjected to analysis of variance. Statistical data from the experiment were inconclusive. Though subjects clearly identified with members of their own race, this difference in reference groups had no significant effect on interaction between race of subjects and narrator when other criteria were involved.

67. Findley, Warren G. and Bryan, Miriam M. Ability Grouping: Do's and Don't's. Integrated Education, 9(5): 31-36, September-October 1971.

This article contains the major findings and recommendations of a study of ability grouping. Among the findings were: virtually all ability grouping plans depend on tests of aptitude or achievement as an integral feature "and socioeconomic and social class differences are increased by streaming; reduced by non-streaming."

68. Findley, Warren G. and Bryan, Miriam M. Ability Grouping: 1970--Status, Impact, and Alternatives. Athens, Ga.: University of Georgia, Center for Educational Improvement, 1971. 98p. ED 060 595 (MF-$0.65; HC-$3.29).

(cont.)

The examination of the use of ability grouping of students begins with presentation of the questionnaire responses from 328 school districts concerning how and how much ability grouping is practiced within their systems, on what basis students are assigned to groups, and how many poor or non-white students are involved. Following is a summary of research relevant to the impact of ability grouping on school achievement, affective development, ethnic separation, and socioeconomic separation. Consideration of the problems and utilities involved in the use of tests for grouping children with limited backgrounds focuses on test reliability and validity, cultural bias, publishers test information, and use of tests with disadvantaged and Mexican American groups. The final section contains a series of brief accounts of alternative strategies to ability grouping.

69. Findley, Warren G. and Bryan, Miriam M. Ability Grouping: 1970--III. The Problems and Utilities Involved in the Use of Tests for Grouping Children with Limited Backgrounds, and Alternative Strategies to Such Grouping. Athens, Ga., University of Georgia, College of Education, 1970. 56p. ED 048 383 (MF-$0.65; HC-Available from Dr. Morrill M. Hall, University of Georgia, Athens, Ga. 30601).

Problems in the interpretation of standardized tests used to group children of limited background, cultural bias in tests, and the misuse of tests are considered. Reports on the use of specific tests with disadvantaged students are reviewed and some of the efforts being made to provide better interpretive data are discussed. Alternative strategies to homogeneous and heterogeneous ability grouping are suggested and described in some detail. The mutually compatible strategies include: individualized instruction, stratified heterogeneous grouping, student tutoring, team teaching, and early childhood education. An extensive bibliography and list of test references are provided.

70. Fishman, Joshua et al. Guidelines for Testing Minority Group Children. Washington, D. C., Society for the Psychological Study of Social Issues, 1963. 27p. ED 001 649 (MF-$0.65; HC-$3.29).

Standardized tests currently in use present three principal difficulties when used with disadvantaged minority groups. Many characteristics of minority group children affect test performance. The lower-class child will tend to be less verbal, less self-confident, less motivated toward academic achievement, less competitive intellectually, less exposed to stimulating materials in the home, less knowledgeable about the world, and more fearful of strangers than the middle class child. The second difficulty is that the significance of the tests for predictive purposes may be quite different for the minority groups

than for the standardization and validation groups. Factors that affect the test scores but which may have little relation to the criterion, such as test-taking skills, anxiety and motivation, may impair predictive validity. Also, the criteria which a test is predicting are usually more complex than the test itself. It is important to recognize the influence of such other factors as personality and background, which may be related to criterion performance. Test results cannot reveal the degree to which the status of disadvantaged children might change if environmental opportunities and incentives for learning were improved. Guidance and special training are therefore very important. Tests labeled "culturally unfair" may be valid predictors for school criteria which may be socially unfair. Most "culture free" tests have low predictive validity for academic work.

71. Fitzgibbon, Thomas J. Evaluation in the Inner City. New York, N. Y., Harcourt Brace Jovanovich, Inc., 1970. 148p.

The invitational conference reported in this work examined some of the problems of measurement and evaluation in the inner city. Both the traditional tools and techniques of measurement in evaluation are being questioned. Thus, a need was seen for clearly justifying the existence of testing and research efforts and enlisting the support of the community in these efforts. The papers treated such topics as the bias of testing, its impact on the child's self-image and on the school system itself, the role of racism in fostering anti-research attitudes, and parents' reactions to testing. Those who participated in the conference included experts in research and measurement, parents and school representatives from the Philadelphia community, and representatives of other school systems. The discussions which followed each lecture are included.

72. Fitzgibbon, Thomas J. The Use of Standardized Instruments with Urban and Minority-Group Pupils. New York, Harcourt Brace Jovanovich, Inc., 1971. 9p. ED 068 505 (MF-$0.65; HC-$3.29).

The position of the Harcourt Brace Jovanovich Test Department is set forth for the use of standardized instruments with urban and minority group pupils. Concern is with the effectiveness and usability of the instruments by the organization that published them. Renewed and intensified attention is given to certain aspects of the traditional test-development enterprise to ensure the appropriateness of the instruments in view of concern and controversy about testing in relation to various social issues. Some examples are given of specific ways in which the test development procedures have been modified in relation to the matter of content validity. Other areas of concern include appropriate norms as a national frame of reference

and testing in written or dictated English pupils whose native language is not English. It is felt that the responsibility of test publishers includes the consideration of test interpretation and use of test results, and efforts in these areas are being expanded. A short list is included of additional publications concerned with the question of effective communication of standardized test information to the community.

73. Flaugher, Ronald L. Minority Versus Majority Group Performance on an Aptitude Test Battery. Project Access Research Report No. 3. Princeton, N. J., Educational Testing Service, 1971. 23p. ED 056 081 (MF-$0.65; HC-$3.29).

Black students from two cities and Mexican-American students, both male and female, showed small but consistent tendencies to perform better, relative to White groups, on three nontraditional measures: tests of inductive reasoning, spatial scanning, and associative memory. These measures showed somewhat less discrepancy between the groups than did tests of the more traditional verbal and mathematical aptitudes. The results seemed compatible with the suggestions of the Commission on Tests to expand the number of measures included in traditional testing programs. Other measures of these same aptitudes, and their validities, remain to be explored in this context.

74. Flaugher, Ronald L. Project Access Research Report No. 2: Patterns of Test Performance by High School Students of Four Ethnic Identities. Princeton, N. J., Educational Testing Service, 1971. 33p. ED 055 102 (MF-$0.65; HC-$3.29).

Among the four participating ethnic groups in Project Access for Los Angeles (black, white, Mexican-American, and Oriental), the patterns of test performance that emerged were surprisingly similar to those from two previous studies of first-grade children, even though Access involved eleventh graders. Partially replicated was the finding that overall level of performance varied with social class, while the patterns remained rather constant within ethnic group, regardless of social class. This new evidence adds to the strength of the argument that our educational systems should be attending to, and utilizing, the variety of talents and aptitudes presented by the students, rather than attempting to treat everyone as a member of the majority group.

75. Flaugher, Ronald L. Testing Practices, Minority Groups, and Higher Education: A Review and Discussion of the Research. Princeton, N. J., Educational Testing Service, 1970. 36p. ED 063 324 (MF-$0.65; HC-$3.29).

Some of the controversial issues involved in the use of objective tests by institutions of higher education, as this use affects the selection and attendance by members of minority groups, are reviewed. Admissions committees now rely on the ability of a test to predict students' performance at their institution to guide their selection. However, minority group members have criticized such uses of test scores. Three potential sources of bias against minority groups include: irrelevance of the test content, particularly verbal content, to their culture and background; discriminatory administration of the testing program; and discriminatory use of the test results. Research investigating the comparative performance of minority and majority group members, the predictive validity of tests, and the influence of the testing environment on performance is reviewed. Efforts to isolate culturally biased test items have been unsuccessful. Evidence indicates that minority group members tend to score less well on most tests; however, tests seem to validly predict academic success regardless of the student's background. In addition, the physical and psychological atmosphere in which the test is administered seems to have a significant influence on performance. Suggestions for supplementary research are delineated.

76. Fleming, Elyse S. and Anttonen, Ralph G. Teacher Expectancy or My Fair Lady. American Educational Research Journal, 8(2): 241-252, March 1971.

"It appears that, in the real world of the teacher using IQ test information, the self-fulfilling prophecy does not operate as Rosenthal [Pygmalion in the Classroom, 1968] hypothesizes.... The present study suggests that teachers assess children, reject discrepant information, and operate on the basis of previously developed attitudes toward and knowledge about children and tests." (pp. 250-251)

77. Ford, Nick Aaron and Turpin, Waters E. Improving the Reading and Writing Skills of Culturally Disadvantaged College Freshmen. Final Report. Baltimore, Md., Morgan State College, 1967. ED 019 367 (MF-$0.65; HC-$3.29).

One of the conclusions was that "there is evidence that available 'standardized tests' do not reveal accurately the intellective potential of culturally disadvantaged student." (p. 34)

78. Fowles, Barbara R. and Kimple, James A., Jr. Language Tests and the "Disadvantaged" Reader. Reading World, 11(3): 183-195, March 1972.

This paper examined three widely used standard tests designed to measure various aspects of linguistic skill: the Wepman Test of Auditory Discrimination, the Harrison-Stroud Reading Readiness Profile and the Illinois Test of Psycholinguistic Ability. It was found that all three favor the white middle class child over the minority child.

79. Freeberg, Norman E. Assessment of Disadvantaged Adolescents: a Different Approach to Research and Evaluation Measures. Journal of Educational Psychology, 61(3): 229-240, June 1970.

This study presents an attempt to overcome a number of the defects customarily attributed to conventional measures by the construction and application of a test battery tailored specifically in content, format, and administration to disadvantaged adolescent groups.

80. Freeberg, Norman E. Development of Evaluation Measures for Use with Neighborhood Youth Corps Enrollees. Final Report. Princeton, N. J., Educational Testing Service, 1968. 143p. ED 051 288 (MF-$0.65; HC-$6.58).

The paper-and-pencil instruments devised attempt to compensate for deficiencies in currently available tests by shaping content, format, and administration to the academic level and cultural background of the intended respondents.

81. Fremer, J. Criterion-Referenced Interpretations of Survey Achievement Tests. Princeton, N. J., Educational Testing Service, 1972. 35p. ED 065 533 (MF-$0.65; HC-$3.29).

This paper discusses criterion-referencing as a problem of validating tests for particular inferences about human behavior. Presented are several methods with their limitations for validating tests for making inferences to a particular criterion or to several criteria of interest.

82. Friedenberg, E. F. Social Consequences of Educational Measurement. In: Proceedings of the 1969 Invitational Conference on Testing Problems. Princeton, N. J., Educational Testing Service, 1969, pp. 23-30.

This paper presents two functions of educational measurement: 1) assessment of individual competence and 2) the assignment of persons with certain test scores to idealogically acceptable social categories. The author discusses why the latter function is the more important and discussion is presented on some of the social consequences that may result from this function and the groups these consequences have the greatest effect on.

83. Gardner, John W. Excellence: Can We Be Equal and Excellent Too? New York, N. Y., Harper & Row, 1961. 171p.

It is asserted that the easiest and laziest thing to do is to sort youngsters out by aptitude scores and forget the rest. The sorting of individuals in a society is an exceedingly serious and explosive business. Because the consequences for the individual are so serious, the final weighing of the evidence must be made by a qualified and responsible human being rather than a machine.

84. Garfunkel, Frank and Blatt, Burton. The Standardization of Intelligence Tests on Southern Negro School Children. Training School Bulletin, 60: 94-99, August 1963.

"The criteria for success and the means for attaining it must be reevaluated at the same time that predictive instruments are developed." (p. 94)

85. Gentile, J. R. Toward an Experimental Analysis of Reasoning on the Scholastic Aptitude Test: A Pilot Study. Research Memorandum 66-26. Princeton, N. J., Educational Testing Service, 1966.

This study presents an outline of an analysis of the verbal reasoning process in terms of the difficulties that culturally deprived students encounter when they are presented with typical SAT items.

86. Getsinger, Stephen H. et al. Self-Esteem Measures and Cultural Disadvantagement. Journal of Consulting and Clinical Psychology, 38(1): 149, February 1972.

Metropolitan sixth-grade students in St. Louis of ages 11-14 from various SES levels were studied. The instruments administered were the Coopersmith Self-Esteem Inventory, the Soares and Soares Test, and the Ziller Test. The findings indicate either that race and SES are inconsequentially related to self-concept in the age group studied, or that typical self-concept measures are not sensitive to differences that exist.

87. Gimmett, Sadie A. Situational Tests for Evaluation of Intervention Programs: A Position. Tucson, Ariz., Arizona University, 1970. 29p. ED 049 308 (MF-$0.65; HC-$3.29).

Lower verbal ability and interpretation difficulties of poor reduce the effectiveness of achievement tests for assessing the results of experimental education. The advantages of situational tests are noted and it is suggested that tests of this type are necessary for more adequate representation of the effects of intervention, given the characteristics of the target population.

88. Goldieri, Anthony A. et al. The Effects of Verbal Approval Upon the Performance of Middle- and Lower-Class Third-Grade Children on the WISC. Psychology in the Schools, 9(4): 404-408, October 1972.

Results of the study indicate that regardless of the presence or absence of verbal approval by the examiner during administration, there was no significant difference in children's test performance on the WISC.

89. Goodman, Kenneth S. Language Difference and the Ethno-Centric Researcher. 1969. 9p. ED 030 107 (MF-$0.65; HC-$3.29).

The author lists steps for constructing a research study guaranteeing "statistically significant results when comparing two populations which differ linguistically." Many of these steps are direct quotations from typical research reports--choose a control group as much like yourself as possible; assume your own dialect is standard; encode all directions, questions, and answers in your own dialect; judge responses as correct only if they are properly stated in your dialect; use experiences drawn from the control (your own)

group; follow a rule of thumb: if something is important to you, it is important; judge all data as deviation from the control group; you are the norm, all else is deficiency. This facetious model for research is based on a " total fiction: that language difference and language deficiency are synonymous." Every child achieves a basic mastery of his dialect well before beginning school and can express anything important to him to the people in his speech community. He speaks his dialect grammatically: if he didn't, he couldn't be understood since grammar is the system of language and all language is systematic. Because "all people are ethno-centric," we build our stereotypes out of differences between ourselves and others. We need objective humility for effective research, not "elitist" views.

90. Gordon, Edmund W., ed. Access and Appraisal: Continuing Education, Higher Education, Career Entry. The Report of the Commission on Tests to the College Entrance Examination Board. New York, N.Y., ERIC Clearinghouse on the Urban Disadvantaged, 1970. 12p. ED 049 324 (MF-$0.65; HC-$3.29).

This paper asserts that considerable modification and improvement are needed if the tests and services of CEEB are to support equitable and efficient access to the system of mass postsecondary education.

91. Gordon, Edmund W. Problems in the Determination of Educability in Populations with Differential Characteristics. 1969. 34p. ED 033 174 (MF-$0.65; HC-$3.29).

"When the characteristics of the target population vary greatly and traditional approaches to education prove ineffective, a shift in educational appraisal may be necessary from quantitative measurement and static prediction to qualitative measurement and dynamic prescription." (p. 25)

92. Goslin, David A. Criticism of Standardized Tests and Testing. New York, N.Y., College Entrance Examination Board, 1967. 34p. ED 039 392 (MF-$0.65; HC-$3.29).

The concluding remarks of this paper focus on the consideration of the social effects of tests, and on the need to develop rational and systematic policies regarding use of tests with the culturally disadvantaged, dissemination of test results, and the problem of invasion of privacy.

93. Goslin, David A. The Search for Ability; Standardized Testing in Social Perspective. New York, N. Y., Russell Sage Foundation, 1963.

This book is one of a series of reports on the social consequences of ability testing by the Russell Sage Foundation. It analyzes the change which has taken place in our society during the past half-centuryin the development and use of ability tests in assigning individuals to positions in society, and in creating opportunities for social mobility. Among the areas covered are testing in business, industry, the government and the military, prediction, and the impact of ability testing on individuals or groups of individuals.

94. Goslin, David A. The Social Consequences of Predictive Testing in Education. Revised Draft. 1965. 22p. ED 018 219 (MF-$0.65; HC-$3.29).

"Clearly valid criticisms of various uses of test scores (for example, the employment of cut-off scores or the rigid use of tests with culturally deprived groups) may be raised in the light of our knowledge that tests are far from being precise instruments. But I do wish to suggest that we have come far enough to pose an additional set of questions about tests--those related to the effects of testing, regardless of validity, and especially including those cases in which the test adequately performs the task expected of it.... A predictive test will typically result in more attention for the child who does well on the tests than for the child who does poorly." (pp. 2-3)

95. Goslin, David A. The Social Impact of Testing. Personnel and Guidance Journal, 45: 679-682, March 1967.

Attention should be directed at the social consequences of continued widespread use of tests in terms of their impact on the individuals involved and the groups that use them.

96. Goslin, David A. Standardized Ability Tests and Testing. Science, 150: 851-855, 1968.

This article discusses some of the major issues that are raised in connection with ability tests and testing in general. The validity of current criticisms of test is discussed. Among the conclusions is that an effort should be made to develop national and systematic policies on 1) the use of tests with the culturally disadvantaged, 2) the dissemination of test results, and 3) the problem of invasion of privacy.

97. Goslin, David A. Teachers and Testing. New York, N. Y., Russell Sage Foundation, 1967.

This report deals with the role of the teacher in standardized testing. The administration of tests by teachers, opinions of the tests held by teachers, the use teachers make of the tests are discussed.

98. Green, Donald Ross. Biased Tests. Monterey, Calif., CTB/McGraw Hill, 1971. 10p. ED 054 208 (MF-$0.65; HC-$3.29).

This paper is concerned with the accusations made by such groups as the Association of Black Psychologists in their call for a moratorium on testing, that standardized tests are biased. A biased test measures one trait on one group of people but a different trait in a second group. Evidence about the amount of bias in tests is thin. Bias must be determined by research on each instrument. A commitment to such research is in order. If bias is found, reasonable courses of action include test revision, alteration in interpretation, and discontinuance of testing.

99. Green, Donald Ross. Racial and Ethnic Bias in Test Construction. Final Report. Monterey, Calif., CTB/McGraw Hill, 1971. 104p. ED 056 090 (MF-$0.65; HC-$6.58).

To determine if tryout samples typically used for item selection contribute to test bias against minority groups, item analyses were made of the California Achievement Tests using seven subgroups of the standardization sample: Northern White Suburban, Northern Black Urban, Southern White Suburban, Southern Black Rural, Southern White Rural, Southwestern Mexican Urban and Southwestern Anglo-American Suburban. The best half of the items in each test were selected for each group. Typically about 30% of the items in the upper half of the distribution of item-test correlations for a group on a test did not meet this criterion with another group. By this criterion minority groups were relatively similar as were the three suburban groups. The resulting unique item tests did not correlate well with each other. Scores of minority groups were relatively better on the selected items. Thus, standard item selection procedures produce tests best suited to groups like the majority of the tryout sample and are therefore biased against other goups to some degree. This degree varies. Ways to minimize this bias need to be developed.

100. Griffiths, Anita. Tests as a Form of Rejection of Minority Group Children. Paper presented at Canadian Guidance and Counseling Assn. Convention, Toronto, 1971. 8p. ED 055 289 (MF-$0.65; HC-$3.29).

The author claims that the dignity and self-esteem of minority group children in particular are effectively destroyed in schools. Racial and ethnic prejudices create the visious cycle of low expectations--lower espectations. Abilities are further undermined by the emphasis in schools on verbal skills and reading as opposed to non-verbal skills. Ability tests, standardized with middle class white children and geared to their value system, are used to classify all children for the rest of their lives; however, most minority group members cannot do well because of the test's irrelevance or senselessness. Because these children's backgrounds and experiences are too diverse, standardized tests provide little help in achieving the goal of equal educational opportunity when used for grouping children. However, if they are used to develop recommendations for individualized instruction, to provide experiences appropriate to the development of each child, they have value. Several examples of intelligent test items which are potentially unintelligible or likely to be misinterpreted, are cited.

101. Guthrie, George M. et al. A Culture-Fair Information Test of Intelligence. University Park, Pa., Pennsylvania State University, 1970. 11p. ED 054 212 (MF-$0.65; HC-$3.29).

An intelligance test for use with rural Filipinos was developed using 25 general information questions. The questions bear on experiences common to most rural Filipinos. The methods of development and analyses are described. It is suggested that this strategy could be used with other peasant groups where an assessment of intelligence is needed in order to rank members of one or adjacent communities. This technique does not enable one to compare groups from markedly different cultural backgrounds.

102. Harris, Major L. Testing and Evaluation in Higher Education and Its Effect on Racial Minorities. 1971. 9p. ED 053 213 (MF-$0.65; HC-$3.29).

Various studies have illustrated the effect of environment upon the academic achievement of youth and adults from racial minorities, especially those who reside in poverty communities. Educators who depend upon so-called measures of academic aptitude perform a disservice to students from racial minorities. Not only do these students become educationally crippled by this kind of single-factor decision making, but society also loses the talents of individuals

whose potentials are too complex to accurately determine by scores obtained on tests of academic aptitude and achievement. Students of racial minorities could benefit more from diagnostic examinations which will assist them in analyzing and eliminating their learning difficulties.

103. Hayman, John L., Jr. and Colendine, Jersy. Survey of Testing in the Great City Schools. Washington, D. C., Council of the Great City Schools, 1970. 101p. ED 055 122 (MF-$0.65; HC-$6.58).

Recognizing 1) the problems which exist in testing children of different cultural backgrounds, and 2) the issue of the use of these test results, this study was undertaken to determine to what extent the large urban school systems, particularly the Great Cities, have been/are/will be involved in test instrument standardization and to sample several aspects of test results. Extensive lists of the tests in use, where, and with what programs make up the major portion of the report.

104. Health Services and Mental Health Administration, Division of Emergency Health Services (DHEW), Rockville, Md. School Achievement of Children by Demographic and Socioeconomic Factors, United States. 1971. 78p. ED 065 799 (MF-$0.65; HC-$3.29).

This is the second report on the school achievement of children 6-11 years old in the noninstitutionalized population of the United States as estimated from the Reading and Arithmetic subtest data of the Wide Range Achievement Test (WRAT) obtained in the Health Examination Survey of 1963-65. It contains findings by selected demographic and socioeconomic variables. Three different survey programs are utilized in the National Health Survey. The report contains findings from the Reading and Arithmetic subtests of the WRAT by age, sex, race, region, size of place of residence and population change from 1950 to 1960, family income, education of parent, and grade in school. These findings are shown in terms of raw scores, standard scores, and grade equivalents. A battery of tests, designed to assess some aspects of mental growth and behavioral development, was included as part of the standardized examination. Significant differentials in school achievement were found over all variables which were considered.

105. Hellmuth, Jerome, ed. Compensatory Education: A National Debate. Disadvantaged Child, Volume 3. New York, N. Y., Brunner/Mazel, 1970.

Section II of this book deals with "Testing and Evaluation." Among the articles are: "Another Look at Culture-fair Testing" by Arthur Jensen, and "Has Evaluation Failed Compensatory Education?" by Herbert Zimiles.

106. Hess, Robert D. et al. Techniques for Assessing Cognitive and Social Abilities of Children and Parents in Project Head Start. Chicago, Ill., Univ. of Chicago, 1966. 139p. ED 015 772 (MF-$0.65; HC-$6.58).

In order to develop and evaluate measuring instruments for assessing the cognitive capacities of disadvantaged children, extensive testing of preschool pupils at 4 headstart centers was conducted. It was hoped that as a result of this comprehensive testing and test-instrument evaluation, a battery of effective instruments could be identified that would reliably predict preschool children's subsequent school achievement, evaluate their school readiness, and point out areas of special disability.

107. Hewer, V. H. Are Tests Fair to College Students from Homes with Low Socioeconomic Status? Personnel and Guidance Journal, 43(8): 764-769, April 1965.

The conclusions of this study were: a) college grades can probably be predicted more efficiently by social group than for the total group, (b) grades for students from homes of superior cultural advantage are not over-predicted nor are grades under-predicted for those of lesser cultural advantage; c) there is no consistent relationship between social origin and achievement of college grades.

108. Hills, John R. and Stanley, Julian C. Easier Test Improves Prediction of Black Students' College Grades. Journal of Negro Education, 39(4): 320-324, Fall 1970.

"The two subtests of level 4 of the School and College Ability Tests (SCAT), for school grades 6-8 are shown to predict freshman-year grades in the three predominantly Negro coeducational four-year colleges of a Southern state considerably better than did the Scholastic Aptitude Test (SAT)..." (p. 324).

109. Hoepfner, Ralph and Strickland, Guy P. Investigating Test Bias. Los Angelos, Calif., University of California, Center for the Study of Evaluation, 1972. 35p. ED 066 443 (MF-$0.65; HC-$3.29).

This study investigates the question of test bias to develop an index of the appropriateness of a test to a particular socioeconomic or racial-ethnic group. Bias is defined as an item by race interaction in an analysis-of-variance design. The sample of 172 third graders at two integrated schools in a large California school district, included 26 white students, 20 Blacks, 64 Mexican-Americans, and 37 Orientals. In order to obtain the initial information about item by race interaction, the Stanford Achievement Test, Paragraph Meaning subtest was used. Item regression data for six racial pairings were inspected: Whites/Blacks; Whites/Mexican-Americans; Whites/Orientals; Blacks/Mexican-Americans; Black/Orientals; and Mexican-Americans/Orientals. Various methods of establishing the existence and nature of test bias cannot be conclusively demonstrated in a wholly satisfactory manner. One method was nontheless selected and applied to test items administrated to two field-test schools for the purpose of investigating bias. The results of that small-scale study are discussed, but do not offer compelling reasons for the observed racial ethnic differences.

110. Hoffman, B. The Tyranny of Testing. New York, N. Y., Crowell Collier, 1962.

The author presents facts that indicate that objective tests are forming barriers which interfere with the searching for superior talents in the schools, the business community, and in the nation itself. He presents documentation that explains the main effects of the multiple choice tests and he also discusses the limitations and misuse of the I.Q.

111. Hsu, Tse-Chi and Boston, M. Elizabeth. Criterion-Referenced Measurement: An Annotated Bibliography. Pittsburgh, Pa., University of Pittsburgh, Learning Research and Development Center, 1972.

This bibliography is concerned with articles that can fit into the interpretation of criterion-referenced measurement which incorporates absolute standards in measuring human behaviors. It was compiled because of a growing interest in criterion referenced measurement due to the emphasis in recent years in education upon such innovations as programmed learning, individualization of instruction, performance contracting, accountability, and computer testing.

112. Humphreys, Lloyd G. et al. Projection Techniques of Objective Factor Analysis. Champaign, Ill., University of Illinois, 1970. 117p. ED 048 330 (MF-$0.65; HC-$6.58).

This collection of papers, concerned with the nature and theory of intelligence, forms part of a project to integrate test and factor theory with the empirically functional relationships involving standard intelligence tests.

113. Hunt, J. McV. Intelligence and Experience. New York, N. Y., Ronald Press, 1961.

"The validity of tests can be divided into at least two kinds, predictive validity, or the capacity of a score from a testing at one time to predict the score from a testing at a later time in the individual's life, and criterion validity, or the capacity performance of tests to predict performance at the same stage of development in various types of life situations. The fact that scores from repeated testing in infancy, when the rate of change in the intellectual structures is greatest, fluctuate radically and the fact that such scores show poor predictive validity are precisely what would be expected if intellectual capacity depends to a considerable degree upon the child's encounters with his environment, but such fluctuations are highly embarrassing to the assumptions of fixed intelligence and predetermined development." (p. 360)

114. Hurt, Maure Jr. and Mishra, Shitala P. Reliability and Validity of the Metropolitan Achievement Tests for Mexican-American Children. Educational and Psychological Measurement, 30(4): 989-992, Winter 1970.

Reliability and validity of the MAT were found to be high for 186 Mexican-American children who were from elementary schools serving "poverty" areas as defined under Title I of ESEA and who had been exposed to a specific curriculum designed to accelerate language development. It is recommended that caution be used in generalizing the results of this study to children of 1) other cultural groups and 2) those from schools with a more traditional curriculum.

115. Hutchinson, June O'Shields. Reading Tests and Nonstandard Language. Reading Teacher, 25(5): 430-437, February 1972.

The linguistic literature on black dialect since 1966 is briefly reviewed,

the main point being that despite the general agreement that children should be evaluated in terms of his own speech community, most children are still tested with the Metropolitan Reading Test which was not designed or standardized for children who speak a "black dialect." These children do poorly on these tests, self-image is harmed, and they may be doomed to a series of such failures resulting in eventual dropout.

116. Hutton, Jerry B. Practice Effects on Intelligence and School Readiness Tests for Preschool Children. Training School Bulletin, 65(4): 130-134, February 1969.

Scores obtained on three screening tests administered to 68 children during the fourth week of Summer Head Start were analyzed. Only minimal significance ($p < .10$) was obtained when comparing test scores changes according to order of administration (practice effects). Further analysis of the results suggested that both the order and type of test may differentially influence the degree of test-retest change. Implications are drawn for assessing the treatment of effects of preschool programs.

117. Jackson, R. Developing Criterion-Referenced Tests. T.M. Report Number 1. Princeton, N. J., Educational Testing Service, 1971.

The article points out that interest in criterion-referenced tests has risen in recent years and that these measures have been considered particularly desirable in areas where diagnostic information is needed. It is also pointed out that there is some doubt as to whether "pure" criterion-referenced tests can be developed for complex domains.

118. Jacobs, John F. and DeGraaf, Carl A. Expectancy and Race: Their Influences upon the Scoring of Individual Intelligence Tests. Final Report. Carbondale, Ill., Southern Illinois University, Dept. of Special Education, 1972. 46p. ED 068 529 (MF-$0.65; HC-$3.29).

A research study is reported which investigated the influence of the variables of race (examiner and child race) and expectancy (high and low) upon the scoring of individual IQ tests. Video tapes of two children (one black, one white, 10 years old, in the 85 to 95 IQ range) being administered the Wischler Intelligence Scale for Children were shown to 32 practicing school psychologists for their evaluations. Half of the psychologists were led to believe

the children were brighter than they actually were (high expectancy); the other half, that they were duller (low expectancy). The expectancy and order of child presentation (high-black, high-white, low-white, low-black, etc.) were randomized. The results indicate that in the case of expectancy white and black examiners score children (white or black) in a similar manner; i.e., in the direction of expectancy. Expectancy has the more influence upon scores obtained when children and examiners are of the same race. The implications for psychological examiners are that expectancy of student ability has a predictable influence upon the IQ attributed to that student. The question remains, however, as to whether such a small average difference in IQ scores has any educational significance.

119. Jensen, Arthur R. How Much Can We Boost IQ and Scholastic Achievement? Harvard Educational Review, 39(1): 1-123, 1969.

The author begins by questioning a central notion upon which compensatory education efforts and other educational programs have been based: that IQ differences are almost entirely a result of environmental differences and the cultural bias of IQ tests. After tracing the history of IQ tests, Jensen carefully defines the concept of IQ, pointing out that it appears as a common factor in all tests that have been devised thus far to tap higher mental processes. Jensen further argues that "although the IQ is certainly not constant, it seems safe to say that under normal environmental conditions it is at least as stable as developmental characteristics of a strictly physical nature." (p. 19)

120. Jensen, Arthur R. Improving the Assessment of Intelligence. Paper presented at the Western Psychological Assn. Annual Meeting, Portland, Ore., 1972. 9p. ED 062 483 (MF-$0.65; HC-$3.29).

Contrary to popular opinion, it is very difficult to find any objective evidence of culture bias that could account for social class and racial differences in performance on current standard tests of intelligence, even those like the Peabody Picture Vocabulary Test (PPVT), which give the appearance of being highly culture-loaded. They may be culture-loaded, but there is no evidence one has been able to find that the culture-loading differentially affects the performance of Negro and white children. Difference in mean score cannot be a criterion of culture bias. One must seek other evidence. The following types of evidence have been examined in the PPVT, the Lorge-Thorndike Intelligence Test, and Raven's Progressive Matrices, these studies having involved very large samples of Negro and white children in several California school districts: (1) The rank order of item difficulty is virtually the same for Negroes

and whites; (2) The matrix of item intercorrelations and the factor structure of these tests is not significantly different for white and Negro samples when these are roughly matched for mental age or total score; (3) In multiple-choice tests, there is no systematic or significant racial difference in the choice of distractors on those items that are answered "wrong"; and, (4) The intelligence tests show essentially the same correlations with scholastic achievement in Negro and white samples.

121. Jensen, Arthur R. Intelligence, Learning Ability, and Socio-economic Status. 1968. 21p. ED 023 725 (MF-$0.65; HC-$3.29).

Discussed are the theoretical explanations of the observation that low intelligence quotient (IQ), low socioeconomic status children appear to be brighter in certain ways than low IQ middle class youngsters. The two different theories on IQ as a function of socioeconomic status--environmental or cultural vs. genetically determined biological potential factors--are evaluated. Also presented is a discussion of the importance of cultural bias in tests and of the various correlations of IQ and learning tests.

122. Jensen, Arthur R. Social Class and Verbal Learning. In: Social Class, Race, and Psychological Development edited by M. Deutsch, and others. New York, N. Y., Holt, Rinehart and Winston, 1968. pp. 115-174.

"The group of children identified by intelligence tests and by various criteria of school achievement as 'slow-learners' or 'mentally retarded' actually constitutes a psychologically very heterogeneous group. A necessary adjunct to research on the role of cultural factors in the development of learning abilities is some means of diagnostically differentiating retardation due to environmental factors from retardation due to more basic biological factors.... The standard tests of intelligence are not suitable for this purpose unless we know quite precisely what opportunities for learning the child's past environment has afforded. What we need is some means of independently assessing what might be called basic learning ability--the rate at which new learning can take place, new learning sets can be developed, and so on, as distinct from the amount of learning that has already occurred in the subject's past." (p. 167)

123. Jensen, Philip K. et al. Evaluating Compensatory Education: A Case Study. Education and Urban Society, 4(2): 211-233, February 1972.

Using standardized tests only, the experimental subjects of a compensatory education program showed little or no gain over control subjects. But when later academic grades were compared, the experimental students outperformed the control students.

124. Johns, Daniel Jay. Correlates of Academic Success in a Predominantly Black, Open-door, Public, Urban Community College. 1970. 113p. ED 039 877 (MF-$0.65; HC-$6.58).

Test batteries that predict first quarter freshman grade point averages (GPA's) at an open-door, predominantly black, urban, community college are reported. Entering freshman performance on an academic aptitude test--College Qualification Tests (QT), a reading skills test--Davis Reading Test, and a study habits and attitudes test--Survey of Study Habits and Attitudes (SSHA) was compared with their first-quarter GPA's. Results indicate that performance prediction was relatively greater for male students and that prediction validity was: 1) rather limited using the SSHA, 2) adequate for male students using the CQT variables in multiple correlations. The author finds not only counseling and placement uses--in comparable environments--for these and similar tests, but also suggests their use as a basis on which to restructure admission policies to permit enrollment of a significantly larger proportion of black students.

125. Johnson, Dale L. and Johnson, Carmen A. Comparison of Four Intelligence Tests Used with Culturally Disadvantaged Children, Psychological Reports, 28(1): 209-210, February 1971.

In light of the need for instruments which can be used to evaluate intellectual functioning of disadvantaged preschoolers, responses to 3 brief intelligence tests were compared with Stanford-Binet scores from 29 Head Start children. The Slosson Intelligence Test was judged the best brief substitute for the S-B in that it correlated .79 with the S-B and its mean IQ was virtually the same as the S-B mean. The other 2 tests were the Peabody Picture Vocabulary Test and the Draw-a-Man Test.

126. Johnson, Geraldine F. Metropolitan Tests: Inappropriate for ESEA Pupils. Integrated Education, 9(6): 22-26, November-December 1971.

An analysis of the Metropolitan Tests reveals that they contain symbols not equally familiar and motivating to all socioeconomic groups; they are fixed on verbal symbols, paragraph content, and problem solving most familiar to higher and middle socioeconomic groups.

127. Johnston, William Elbert, Jr. and Scales, Eldridge E. Counseling the Disadvantaged Student. Research Studies Series. 1968. 120p. ED 034 807 (MF-$0.65; HC-Not available from EDRS).

"The problem of rapport between the counselor and the student in the testing situation is difficult enough but the problem may loom even larger when the counselor is white and the student is a Negro. Sincerity, warmth, and the incorporation of other rapport-building attitudes and techniques into the interpersonal contact may be met with protective suspicion, wariness, hostility, or withdrawal behavior." (p. 42)

128. Jones, J. L. Assessing the Academic Achievement of Negro Students. Clearinghouse, 39(2): 108-112, 1964.

This article gives an analysis of the problems that are encountered in the assessment of the academic achievement of black students. It gives certain suggestions that may resolve or correct some of the complex and interrelated problems that are presented.

129. Jones, W. Paul and DeBlassie, Richard R. Social Class Contrasts in Short Term Predictability of Grade 7 Achievement. Journal of Educational Research, 65(1): 11-14, September 1971.

Seventh grade students were used to determine (1) if significant differences existed in mean scores on the Short Test of Educational Ability (STEA) among different social class groups and (2) if there were significant differences in the relationship between STEA scores and academic performance among different social class groups. Results indicated that scores on the STEA distribute themselves along a social class dimension and that significant differences existed in the mean scores of the four social class groups. No significant differences in the predictive relationship of STEA scores and academic performance were indicated whether achievement test scores or academic grades were used as the criteria.

130. Justman, Joseph. Assessing the Intelligence of Disadvantaged Children. Education, 87(6): 354-362, February 1967.

Up to 1967 culture-free tests had been unsuccessful. A culture-free test is defined as "a test in which differences in experiential background and motivation have little or no effect on test scores."

131. Kendrick, S. A. and Thomas, C. L. Transition from School to College. Review of Educational Research, 40(1): 151-179, February 1970.

A review of the literature on the identification and selection of disadvantaged students for admission to college is presented. This literature is primarily an examination of the validity of two traditional predictors of college success: high school scholarship and preadmissions test scores.

132. Kimball, Ronald L. et al. Black and White Vocational Interests on Holland's Self Directed Search (SDS). College Park, Md., University of Maryland, Cultural Study Center, 1971. 11p. ED 065 523 (MF-$0.65; HC-$3.29).

There was some concern that Holland's Self Directed Search for Educational and Vocational Planning (SDS) was inappropriate to use with blacks since it may steer them toward more Realistic occupations not requiring a college education. The purpose of the study was to compare the pattern of vocational planning choices for black and white students and to determine student satisfaction with SDS results. The SDS was administered to 143 blacks during freshman orientation and their results were compared to a random sample of whites. Results indicated that blacks tended to choose Social occupations more often, compared to more Realistic and Investigative choices by whites. That blacks tend to seek social service occupations has been demonstrated in several other studies. There were no differences in satisfaction with SDS results between blacks and whites. Implications for counseling blacks were discussed and the study concluded that the SDS was equally appropriate for blacks and whites.

133. Labov, William. A Study of Non-Standard English. Washington, D. C., Center for Applied Linguistics, 1969. 75p. ED 024 053 (MF-$0.65; HC-$3.29).

"It should be immediately apparent that none of the standard tests will come anywhere near measuring Leon's [a Black child] verbal capacity. On these

tests he will show up as very much the monosyllabic, inept, ignorant, bumbling child of our first interview... We see no connection between verbal skill at the speech events characteristic of the street culture and success in the school-room" (p. 42)

134. Lambert, N. M. The Present Status of the Culture Fair Testing Movement. Psychology in the Schools, 1(3): 318-330, 1964.

The author presents some background in the culture fair testing movement and gives descriptions of some of the culture free tests that exist today.

135. Leary, M. E. Children Who are Tested in an Alien Language: Mentally Retarded? The New Republic, 162(22): 17-18, 1970.

This article discusses English language IQ tests administered to Spanish-speaking children and the subsequent labeling of these children as mentally retarded.

136. Lennon, R. Testimony of Dr. Roger T. Lennon as Expert Witness on Psychological Testing in the Case of Hobson et al. vs. Hansen et al. New York, N. Y., Harcourt, Brace and World, 1967.

The testimony ranges over a wide variety of aspects which include the proper use of various types of standardized tests in a school system; the extent to which tests may be unfair to pupils who are considered to be culturally disadvantaged; the manner in which tests are standardized with special reference to representation in the norm groups of various segments of the population; the issues of predictive validity of tests for various SES and ethnic groups; the question of appropriateness and utility of local versus national norms; the competence of teachers generally in the case of tests and other significant facets of a schools's testing program.

137. Lennon, Roger T. Testing and the Culturally Disadvantaged Child. 1964. 9p. ED 001 430 (MF-$0.65; HC-$3.29).

The use of tests in the education of the culturally deprived is presented. Appropriate use of such tests can make some contributions to the discovery of talent, to the diagnosis of learning difficulties, and to the better guidance

and adjustment of culturally deprived children. Most tests measure ability or aptitude and achievement. The former have received the most criticism in relation to their suitability for the disadvantaged child. Tests of general mental ability are given in order to determine likelihood of success in school work, to inform teachers of performance level and possible rate of progress. Critics claim that there is a built-in bias against lower-class background in these tests. Subsequent misinterpretation results in curtailing of educational opportunities. Tests do use verbal, numberial and academic content that is lacking in the background of disadvantaged children. Interpretation should indicate the child's present state of readiness to cope with classroom learning tasks, not his native potential. The use of achievement tests which show progress toward concrete goals of instruction is not challenged. Suggestions relevant to the culturally deprived are given for the selection, administration and interpretation of such tests.

138. LeSage, W. and Riccio, A. C. Testing the Disadvantaged: An Issue of Our Time. Focus on Guidance, 3(1): 1-7, 1970.

This article explores the validity of the criticisms that are presently argued against testing in relation to the disadvantaged. Possibilities of alternatives to testing are also discussed.

139. Lesser, G. S. et al. Mental Abilities of Children from Different Social-Class and Cultural Groups. Monographs of the Society for Research in Child Development, 30(4), 1965 (Serial No. 102). 233p. ED 003 293 (MF-$0.65; HC-$9.87).

This study examined the patterns among various abilities in six- and seven-year-old children from different social-class and cultural backgrounds. The main intent was to extend the empirical analysis of the development of differential mental abilities in children, but the findings of this research also bear directly upon the problems of building valid and precise assessment instruments for children from different cultural groups.

140. Levine, Murray. Psychological Testing in Children. In: Review of Child Development Research, Volume II, edited by L. W. Hoffman and M. L. Hoffman. New York, N. Y., Russell Sage Foundation, 1966. pp. 257-310.

An overview of the field of psychological testing in children is presented. There is some discussion of attacks on testing on political and social grounds as well as scientific.

141. Linn, Robert L. and Werts, Charles E. Considerations for Studies of Test Bias. Journal of Educational Measurement, 8(1): 1-4, Spring 1971.

This is a technical paper which discusses two problems in the investigation of predictive bias in tests: 1) the effect of unreliability of the predictors and 2) the effect of excluding a predictor from the regression equation on which there are preexisting group differences.

142. Lucas, C. M. Survey of the Literature Relating to the Effects of Cultural Background on Aptitude Test Scores. Research Bulletin 53-13. Princeton, N. J., Educational Testing Service, 1953.

Definitions of some testing terms are presented together with some of the criticism of aptitude tests that have been offered. Discussed are areas such as culturally unbiased tests, methodological shortcomings of comparison studies, racial comparisons, occupational group comparisons, socioeconomic class comparisons.

143. MacArthur, R. S. Mental Abilities in Cross-Cultural Context. 1966. 25p. ED 024 742 (MF-$0.65; HC-$3.29).

Discussed are issues involved in testing the mental abilities of nonWestern, nonurban ethnic groups. Within this context the paper reviews conceptions of intelligence and intellectual potential, prediction under fixed and adaptive conditions, the question of environmental influences, and some formal test factors. Examples are drawn from African populations and Canadian Indian-Metis and Eskimo groups.

144. MacArthur, R. S. and Elley, W. B. The Reduction of Socioeconomic Bias in Intelligence Testing. British Journal of Educational Psychology, 33 (part 2): 107-119, June 1963.

This study was designed to investigate the extent of socioeconomic bias in selected intelligence tests. An attempt was also made to identify and construct instruments which might better measure the intellectual potential of pupils from deprived cultural and educational background. Results show that it is possible to measure a broad component of intellectual ability with significantly less cultural bias than is found in the conventional intelligence test.

145. Mackler, Bernard. Grouping in the Ghetto. Education and Urban Society, 2(1): 80-96, November 1969.

"Some of the dangers in grouping are in the initial selection procedures. How can one determine who is able in kindergarten? What about teacher bias when selections as to ability and behavior are made? What does conforming behavior do to a growing, developing child?" (p. 80)

146. Manning, W. H. The Measurement of Intellectual Capacity and Performance. Journal of Negro Education, 37(3): 258-267, Summer 1968.

"What is perhaps not generally recognized is that widespread application of measurement systems has also led to tests having acquired a societal function, in which the sample of behavior which constitutes the test (and hence scores deriving from it) comes to connote a view of man and the nature of human intellect that has important consequences as these influence social values regarding the worth of the individual and the dignity of various forms of work." (p. 260)

147. Marshall, Patricia. Testing Without Reading. Manpower, 3(5): 7-12, May 1971.

This article discusses the use and validity of vocational aptitude tests with culturally disadvantaged groups.

148. Mathis, Harold J. The Disadvantaged and the Aptitude Barrier. Personnel and Guidance Journal, 47(5): 467-472, January 1969.

The theory and practice of aptitude testing are critically discussed within the context of current efforts to improve the status of the poor.

149. Matuszek, Paula A. and Oakland, Thomas D. A Factor Analysis of General Reading Readiness Measures for Different Socioeconomic and Ethnic Groups. 1972. 9p. ED 068 909 (MF-$0.65; HC-$ 3.29).

First grade students stratified on the basis of their racial-ethnic and socioeconomic status (SES) were studied to determine factors measured by several

readiness (aptitude) tests and to learn whether these factors differed for the children from the various groups. SES was determined by father's occupation, or if absent, mother's occupation; the identification of racial-ethnic backgrounds was left to the teacher. Six groups of children were finally chosen: Middle and lower class Blacks, Mexican-Americans, and Anglo-Americans. The major observation drawn from the data was that factors derived from the tests differed markedly across groups. It was further concluded that present tests were inefficient in that several subtests seemed to measure the same factors. Data also indicated that tests could not be used interchangeably with children from different SES and racial-ethnic backgrounds. (Two tables of factor loadings and a reference list are appended.)

150. Mayer, Martin. The Schools. London, The Bodley Head Ltd., 1961.

Chapter 6 deals with the use of standardized intelligence tests with children from lower socioeconomic backgrounds and the concept of culture fairness. Chapter 17 discusses generally the use of tests and examinations in schools.

151. McDiarmid, G. L. The Hazards of Testing Indian Children. 1971. 21p. ED 055 692 (MF-$0.65; HC-$3.29).

Discusses the role that poverty, health and nutrition, social conflict, language, and test motivation play in relation to interpretation of test data obtained on indian children. The 2 greatest problems reported affecting test reliability, validity, and validity of test interpretation in this context are language and test motivation.

152. Meier, John H. Innovations in Assessing theDisadvantaged Child's Potential. In: Disadvantaged Child, Volume I, edited by Jerome Hellmuth. Seattle, Wash., Special Child Publications, 1967. pp. 173-199.

This paper individually discusses existing tests used with disadvantaged children. An extensive list of references is included.

153. Melville, Robert and Haas, John. Educational Achievement and the Navajo. 1967. 13p. ED 013 172 (MF-$0.65; HC-$3.29).

(cont.)

Subjects in the study included 97 Navajo students and 848 non-Navajo students in instructional program divisions seven through twelve in four secondary schools. Navajo students scored significantly lower on the standardized tests than non-Navajo students on measures of intelligence, self-concept, and motivation, but significantly higher on the anxiety scale. Results of testing in verbal development indicated that Navajo students are operating at a lower level of concept formation than their non-Navajo peers.

154. Mercer, Jane R. Current Retardation Procedures and the Psychological and Social Implications on the Mexican-American. A Position Paper. Albuquerque, N. M., Southwestern Cooperative Educational Lab., 1970. 40p. ED 052 848 (MF-$0.65; HC-$3.29).

In view of the disproportionately large numbers of Mexican American children in special education classes, this position paper presents past and current research findings examining 2 hypothetical causes of the disproportion: discriminatory referral procedures and discriminatory clinical procedures. Examining 1 medium-sized California school system, it was determined that the pupil personnel department's clinical testing procedures rather than referral processes produced the ethnic misrepresentation in special classes. Recommendations are that (1) more refined scales need to be developed for assessing the child's adaptive behavior outside the school and (2) pluralistic norms need to be used in interpreting the meaning of both the IQ and the adaptive behavior score for children of Mexican American heritage. Five tables and 16 references are included.

155. Mercer, Jane R. Sociocultural Factors in the Educational Evaluation of Black and Chicano Children. Presented at the 10th Annual Conference on Civil and Human Rights of Educators and Students, Washington, D. C., 1972. 16p. ED 062 462 (MF-$0.65; HC-$3.29).

Even though the cultural bias of I.Q. tests has been attested to since the 1930's, the I.Q. scores of minority groups are still interpreted without regard for this bias. One result has been to label a disproportionately high number of blacks and Chicanos as mentally retarded.

156. Messick, Samuel and Anderson, Scarvia. Educational Testing, Individual Development, and Social Responsibility. Princeton, N. J., Educational Testing Service, 1970. 24p. ED 047 003 (MF-$0.65; HC-$3.29).

Recent harsh criticisms that educational and psychological tests are unfair and inadequate measures of the capabilities of minority, poverty, and other educationally alienated groups are discussed. The authors suggest that there are two main issues, the first scientific, the second ethical: (1) is a test a valid measure of the characteristics it purports to assess for particular types of individuals in particular circumstances; and (2) the whole question of test use, beginning with whether or not a test should be utilized for a specified purpose. Responsible standards exist for evaluating the adquacy and appropriateness of a test for a particular use, but they are not always applied. The adequacy of measurement and the question of bias, the appropriateness of test use and the question of fairness, the side effects of testing, the problems of misinterpretation and secondary use of tests, the ethics involved, and the social consequences of not testing are other important topics discussed and analyzed in some detail.

157. Messick, Samuel. Personality Measurement and the Ethics of Assessment. American Psychologist, 20: 136-142, 1965.

This article discusses some of the criticisms of psychological testing which the author feels result from conflicting values between the testers and critics and within the field of psychology itself.

158. Milgram, N. A. IQ Constancy in Disadvantaged Negro Children. Psychological Reports, 29(1): 319-326, August 1971.

This paper reports the results of a longitudinal follow-up with conventional intelligence scales of disadvantaged Negro children from age 3 to 8 years. This study analyzed 1) group IQ changes; 2) the change over time for the individual child; and 3) the relationship of IQ to the test-taking behavior of the child as rated along dimensions of cooperation, security, rapport, attention, etc. One of the major findings was that the mean Binet IQ was relatively stable, while the Peabody IQ rose appreciably.

159. Mishro, Shitala P. and Hurt, Maure, Jr. The Use of Metropolitan Readiness Tests with Mexican-American Children. California Journal of Educational Research, 21(4): 182-187, September 1970.

"The results of this study that the tests of the Metropolitan Readiness Tests battery which are verbal in nature have lower reliability and predictive validity than the tests which do not require as much specific ability in English to answer the test items. The potential lack of reliability and validity should be of special concern to those engaged in developmental programs for culturally deprived." (p. 186)

160. Mitchell, B. C. Predictive Validity of the Metropolitan Readiness Tests and the Murphy-Durrell Reading Readiness Analysis for White and Negro Pupils. Educational and Psychological Measurement, 27: 1047-1054, 1967.

"From the data of this study it would appear that the two readiness tests considered, the Metropolitan and the Murphy-Durrell, perform their function as well with Negro pupils as with white pupils, and that the general level of predictive validity is similar for four racial-ethnic groups studied." (p. 1054)

161. Mollenkopf, W. G. A Study of Secondary School Characteristics as Related to Test Scores. Research Bulletin 56-6. Princeton, N. J., Educational Testing Service, 1956.

Results of this study show that the following items are related to aptitude scores. instructional support per pupil, presence or absence of a community library, community size, region, per cent of support from the state, and per cent of fathers who are high school graduates.

162. Moore, Allen B. Selected Standardized Noncognitive Measures and Their Utility for the Disadvantaged. Adult Education, 22(2): 110-121, Winter 1972.

This article concludes that there are few standardized noncognitive measures available for disadvantaged adults and they appear to have limited utility for investigating the effects of basic education on disadvantaged adults.

163. Mour, Stanley I. Evaluation of Reading in the Classroom, Interpreting the Results of Standardized Tests. 1967. 9p. ED 014 384 (MF-$0.65; HC-$3.29).

If it is recognized that the sole justification in giving a test is to obtain a better understanding of the child, factors that influence the child's performance on the test must be considered in the interpretation of test results. A sufficient knowledge of the child's development, his background, his perception of the test situation, and the amount of knowledge and skills he brings into the test situation would help the teacher differentiate and understand the child who does not perform because he cannot and the one who can perform but would not. The teacher should consider every aspect of the test--what it can measure, the test items themselves, the cultural basis of the test, the kind of environment it creates, the statistical aspects, and how the test would contribute to a deeper understanding of the child. While considering all these factors in the interpretation of results, the teacher should remember to interpret to each child the purpose of the test, what the test scores mean, and how these are going to be used.

164. Munday, L. A. Predicting College Grades in Predominantly Negro Colleges. Journal of Educational Measurement, 2: 157-160, 1965.

This article discusses the question as to whether standardized tests have useful validities for predicting relevant criteria for culturally different or disadvantaged students. The author concludes that the grades for socially disadvantaged students are generally as predictable as grades for other students using standardized measurement of academic ability.

165. Murray, Walter J. The I.Q. and Social Class in the Negro Caste. The Southwestern Journal (Langston Univ., Oklahoma), 4(4): 187-201, Fall 1949.

The analysis in this study has been based on the hypothesis that there are statistically significant differences in the performance of Negro children on intelligence tests when these children have been grouped according to social classes. The data obtained by administering 5 intelligence tests to 401 Negro children of this study prove that this hypothesis is true. Some of the difference on performance may be due to the phenomenon of social class as reflected in the test-items themselves.

166. Nalven, Fredric B. et al. The Effects of Subjects' Age, Sex, Race, and Socioeconomic Status on Psychologists' Estimates of the True IQ from WISC Scores. Journal of Clinical Psychology, 25(3): 271-274, July 1969.

The findings of this study are that a child's age or sex does not influence psychologists' judgments as to whether his obtained WISC Full Scale IQ Score represents an accurate estimate of his true potentials. In contrast, a child's social class background, to a great extent, and his race, to a lesser extent, significantly shape psychologists' judgments as to whether the obtained IQ Scores are representative. They suggest that psychologists consider WISC Full Scale IQ scores to be underestimates of children's "True IQ" or "effective intelligence."

167. National Consortia for Bilingual Education, Fort Worth, Texas. Tests in Use in Title VII Bilingual Education Projects. 1971. 188p. ED 050 886 (MF-$0.65; HC-$6.58).

This list of tests representing 72 Title VII projects includes commercially published tests, tests developed by the projects themselves, and tests designed by other entities such as educational laboratories, school districts, and state agencies. Each test is described in terms of the project using the test, language of the project, grade range, test title, origin of the test, stated purposes, age level or grade, level of application, and language of test. It should be noted that "adequacy of measurement, usability, or appropriateness (of the tests) should not be inferred."

168. Nieman, Ronald H. Early Childhood Education Project Evaluation, 1970-71. ESEA Title I. Cincinnati, Ohio, Cincinnati Public Schools, 1971. 17p. ED 060 957 (MF-$0.65; HC-$3.29).

To determine if more organized, purposeful early childhood experiences would enable the child to attain a higher score on standardized tests, in November 1970 the Boehm Test of Basic Concepts was administered to 363 All Day Kindergarten with Pre-School students, 49 All Day Kindergarten without Pre-School students, and 141 Half-Day Kindergarten students. The Metropolitan Reading Readiness Test was administered in April 1971 to a random sample of six All Day Kindergarten classes and six comparable Half-Day Kindergarten classes. A total of 806 kindergarten students, only 509 of whom had pre-school, were administered the Boehm in September 1971. Comparisons were made between pre-school and non-pre-school treatment. The results showed the pre-school

group scored significantly higher than the non-pre-schoolers at the .01 level. Results of the Metropolitan Reading Readiness Test show All Day Kindergarten students to have nearly double the mean of the Half-Day Kindergarten students.

169. Nitko, Anthony J. A Model for Criterion-Referenced Tests Based on Use. Pittsburgh, Pa., University of Pittsburgh, Learning Research and Development Center, 1971. 17p. ED 049 318 (MF-$0.65; HC-$3.29).

"...it is the use to which test results are put that determines their nature and the construction methodology. In instruction, various procedures cannot be considered independently of the instructional context in which they will be used. Particularly important is the integration of test design with instructional design." (p. 9)

170. Oakland, Thomas and Weilert, Evelyn. The Effects of Test-wiseness Materials on Standardized Test Performance of Preschool Disadvantaged Children. 1971. 13p. ED 047 050 (MF-$0.65; HC-not available from EDRS).

This study was an attempt to examine certain variables which apply to test-wiseness factors in assessing preschool and primary grade children. The study identified specific abilities which appear to be prerequisite for taking standardized tests, and to design curricular materials to facilitate the development of these abilities. It was found that the curricular test-wiseness materials had limited value in facilitating performance on the standardized test. However, teachers reported that the majority of children in the experimental group appeared to be more confident on paper and pencil tasks, and to remain task-oriented for longer periods of time.

171. Ozer, Mark N. The Neurological Evaluation of Children in Head Start. In: Disadvantaged Child, Head Start and Early Intervention edited by Jerome Hellmuth. New York, N. Y., Brunner/Mazel, Inc., 1968. pp. 125-136

A neurological examination has been developed which provides a short standardized measure of brain function. Its purpose is to define the extent of the problem of brain disfunction in various populations and to identify the specific child requiring special programming. It has been piloted in a Head Start program and has thus been used to outline the special need of Head Start children.

172. Pasanella, Ann K. et al. Bibliography of Test Criticism. New York, N. Y., College Entrance Examination Board, 1967. 56p. ED 039 395 (MF-$0.65; HC-$3.29).

This is a selected compilation of 47 items relating to criticisms of tests and testing. The items cover the period of ten years immediately preceding the year 1966 and are held to be scholarly writings almost without exception. The annotated listings focus on the following aspects of tests and testing: encouragement of intellectual conformity; erosion on individual freedom of choice; exertion of undue influence on education; invasion of individual privacy; and concealment of true character by masquerading as scientific instruments. The sources of strain in the themes of these books and articles is thought to be three-sided: the tests themselves; the test users; and the test makers.

173. Pascale, Pietro J. and Jakabovic, Shaena. The Impossible Dream: a Culture-Free Test. 1971. 25p. ED 054 217 (MF-$0.65; HC-$3.29).

The study reviewed the formats and psychometric rationale of several alleged culture-fair tests. Advantages and disadvantages of each instrument were examined and implications for compensatory education were discussed.

174. Pasework, Richard A. et al. Relationship of the Wechsler Pre-school and Primary Scale of Intelligence and the Stanford-Binet (L-M) in Lower Class Children. Journal of School Psychology, 9(1): 43-50, 1971.

Using 72 lower class children, concurrent validity of the WPPSI was studied using the Stanford-Binet as the criterion. Correlations between SB and WPPSI were moderately high. Among the WPPSI subtests, vocabulary was significantly lower than all other subtests.

175. Pelosi, John William. A Study of the Effects of Examiner Race, Sex, and Style on Test Responses of Negro Examinees. Ph.D. Thesis. Syracuse, N. Y., Syracuse University, 1968. 209p. (Available from University Microfilms, 300 N. Zeeb Rd., Ann Arbor, Mich. 48106; order no. 69-8642.

This study investigated the influence of examiner race (white versus Negro), style of interaction (warm versus cold), and sex (male versus female) on test

responses of 96 male Negro subjects enrolled in an antipoverty work experience program. The Information Comprehension, Vocabulary, Digit-Symbol, Block Design, and Picture Arrangement subtests of the Wechsler Adult Intelligence Scale were used, together with the Purdue Pegboard and the IPAT Culture Fair Test. Contrary to previous research findings which suggested inadvertent bias due to white examiners, the overall conclusion of this study was that examiner traits did not significantly influence performance by Negro male subjects on seven of the tests. The only exception was the IPAT Culture Fair Test, on which subjects treated coldly performed better under male Negro examiners than female Negroes, and better under white females than white males. Implications and limitations of the study were discussed.

176. Perry, Dallis. Interpreting Standardized Test Scores. St. Paul, Minn., University of Minnesota, Student Counseling Bureau. 1971. 57p. ED 053 201 (MF-$0.65; HC-$3.29).

Principles of test administration, test validity, and accuracy of measurement underlying interpretation of standardized test scores in educational administration, instruction, and guidance are presented.

177. Pfeifer, C. Michael, Jr. and Sedlacek, William E. The Validity of Academic Predictors for Black and White Students at a Predominantly White University. Journal of Educational Measurement, 8(4): 253-261, Winter 1971.

High school grades and both verbal and mathematical scales of the SAT were examined as predictors of college grade point average in groups divided by race and sex. Results indicated that high school grades were not correlated as highly with college grades for black males as for the other 3 groups, although there was no significant difference in the correlation of either SAT with college grades.

178. Phillips, Judith. The Effects of the Examiner and the Testing Situation Upon the Performance of Culturally Deprived Children. Phase I--Intelligence and Language Ability Test Scores as a Function of the Race of the Examiner. Final Report. Nashville, Tenn., George Peabody College for Teachers, 1966. 18p. ED 011 963 (MF-$0.65; HC-$3.29).

Pretest scores obtained prior to language development had indicated that

Negro examiners produced significantly higher Stanford-Binet IQ scores than white examiners. No differences associated with the race and sex of the examiner and child were obtained with 240 Stanford-Binet post-test scores. White examiners produced significantly higher scores than Negro examiners with 160 ITPA post-test results. The greatest difference between Negro and white examiners occurred on the ITPA vocal encoding subtest, where whites produced more spontaneous vocalization. (ITPA: Illinois Test of Psycholinguistic Abilities)

179. Pinard, Adrien and Sharp, Evelyn. I.Q. and a Point of View. Psychology Today, 6(1), June 1972.

The University of Montreal's Institute of Psychology is devising a new intelligence test based on Piaget's theories. It is emphasizing the quality of a child's reasoning rather than quantifying of right and wrong answers. This article is a progress report and describes the construction and interpretation of the test items.

180. Politzer, Robert L. Auditory Discrimination and the "Disadvantaged": Deficit or Difference. English Record, 21(4): 174-179, April 1971.

A major problem in testing the auditory discrimination of disadvantaged youth is their lack of contact with standard dialect, which results in their failure to make categorizations made by standard speakers.

181. Popham, W. James and Husek, T. R. Implications of Criterion-Referenced Measurement. Journal of Educational Measurement, 6(1): 1-9, Spring 1969.

This discussion examines the implications of two approaches to measurement, norm-referenced and criterion-referenced, with respect to variability, item construction, reliability, validity, item analysis, reporting, and interpretation. Criterion-referenced measurement is discussed more fully.

182. Puerto Rican Study. Developing a Program for Testing Puerto Rican Pupils in the New York City Public Schools. New York, N. Y., Board of Education of the City of New York, 1959.

(cont.)

This paper discusses actual tests which were developed by the Puerto Rican Study. The Study took the position that what was being measured was an index of the pupil's current functional level, not a measure of his innate potential or of the cultural level of his ethnic group.

183. Quay, Lorene C. Negro Dialect and Binet Performance in Severely Disadvantaged Black Four-Year-Olds. Child Development, 43(1): 245-250, March 1972.

Findings, confirming the results of an earlier study by the author using a less deprived population, indicated that young black children do not benefit from having the Binet administered in Negro dialect.

184. Rambo, Lewis M. Effects of Experimenter Bias on Attitudes Toward Controversial Issues. Journal of Negro Education, 38(4): 384-394, February 1969.

This study demonstrates the effect of race of test administrator on the responses of experimental groups of white undergraduate subjects on a comprehensive battery of attitudinal measures.

185. Rankin, Richard J. and Henderson, Ronald W. Standardized Tests and the Disadvantaged. Tucson, Ariz., University of Arizona, Arizona Center for Early Childhood Education and National Laboratory on Early Childhood Education, 1969. 13p. ED 034 594 (MF-$0.65; HC-$3.29).

"A major problem in intellectual measurement has always been speculation concerning the impact of item biasing upon the validity of the obtained measurement. The problem is aggravated when the investigator takes measured intelligence to be somehow related to genetic endowment rather than recognizing that it is simply a statistical abstraction sometimes useful for predicting specific achievements... It seems undeniable that individual's skills emerge through interaction with the experiences available and valued in his cultural or social milieu. Therefore, we would expect children reared in settings which differ in significant ways from the culture of the norming sample to display below average performance on standardized tests. This says nothing about their genetic inheritance." (p. 3)

186. Rice, James A. Head Start Screening: Effectiveness of a Teacher-Administered Battery. Perceptual and Motor Skills, 32(12): 675-678, April 1971.

Effective screening is an important issue in the Head Start program. A battery of tests, such as the one described in this study, can provide an objective basis for answering such questions as: Who should be included or excluded? When should screening take place? Should a child be allowed to repeat the program?

187. Robb, Juanita. Syllabus for the Teaching of Reading to Disadvantaged Youth. Monmouth, Ore., Oregon College of Education, 1967. (Available from Training Center for Teachers of Migrant and Disadvantaged, Oregon College of Education, Monmouth, Ore.).

Problems of diagnosis include the interdisciplinary investigation of physical deterrents to learning, language disabilities, and the assessment of the total picture of their potential to learn. Informal inventories are frequently more helpful than the results on standardized instruments.

188. Rosen, Doris B. Employment Testing and Minority Groups, As Reported in the Literature. Key Issues Series No. 6. Ithaca, N. Y., State University of New York, 1970. 35p. ED 048 507 (MF-$0.65; HC-$3.29).

This report reviews the employment testing of minorities and discusses the limitations of the test results. Statistical validation of test results is stressed as a means to avoid cultural bias and avoid discriminatory screening tests.

189. Rosen, Pamela and Horne, Eleanor V. Tests for Spanish-Speaking Children: An Annotated Bibliography. Princeton, N. J., Educational Testing Service, 1971. 14p. ED 056 084 (MF-$0.65; HC-$3.29).

Brief annotations of currently available instruments appropriate for use with Spanish-speaking children are presented. Measures in such areas as intelligence, personality, ability, and achievement are included and separate indexes are given. The annotation provides information concerning the purpose of the test; the groups for which it is intended; test subdivisions or tested skills, behaviors, or competencies; administration; scoring; interpretation; and standardization. An alphabetical listing of the tests which indicates the ages for which each is considered suitable is also included.

190. Rosenberg, Leon A. A Culture-Fair Instrument for Intellectual Assessment. In: Disadvantaged Child, Head Start and Early Intervention edited by Jerome Hellmuth. New York, N. Y., Brunner/Mazel, 1968. pp. 77-92.

This article discussed the Johns Hopkins Perceptual Test as a "culture-free" instrument.

191. Rosenberg, Leon A. et al. The Johns Hopkins Perceptual Test, The Development of a Rapid Intelligence Test for the Pre-school Child. Baltimore, Md., Johns Hopkins University, 1966. 17p. ED 020 787 (MF-$0.65; HC-$3.29).

Statistically significant correlations of the scores on the JHPT with scores on the PPVT, the DAP, and the Columbia Mental Maturity Scale (CMMS) had been obtained. A comparison was made of the performance of 37 upper and 52 lower socioeconomic status children on the JHPT, the PPVT, and the DAP. The differences in mean scores between groups was highest for the PPVT, less marked for the DAP, and statistically insignificant for the JHPT, suggesting that it may be a culture-free instrument.

192. Rosenberg, Leon A. and Stroud, Michael. The Limitations of Brief Intelligence Testing with Young Children. Baltimore, Md., Johns Hopkins University. 13p. ED 020 774 (MF-$0.65; HC-$3.29).

A study examined the validity of two brief intelligence scales, the Peabody Picture Vocabulary Test (PPVT) and the Columbia Mental Maturity Scale (CMMS), to see if these tests overestimated the incidence of intellectual retardation among poverty area children. These two tests and an experimental measure of intelligence, the Johns Hopkins Perceptual Test (JHPT), had been given to 28 kindergarten children. A significantly less retardation incidence was found with the JHPT. Actual percentages were 69.56 on the PPVT, 17.86 on the CMMS, and 3.57 on JHPT. Therefore, as a criterion measure, the Stanford-Binet was administered to each child. It was found that the children tested on a much higher level of intelligence than reported on the PPVT and CMMS, although the CMMS exaggerated less the degree of retardation in the same school. A sample of first grade children representing an entire kindergarten class of the previous year was given the PPVT, CCMS, and the Stanford-Binet. Twenty-nine children participated. It was found that the PPVT still underestimated intelligence to a statistically significant degree. The mean IQ score increased 18.86 points on PPVT performance, indicating the influence of educational ex-

perience. Although the Stanford-Binet still placed all children in the nonretarded category, the PPVT rated 24.14 per cent and the CMMS Scale, 10.34 percent as retarded. A brief culture-fair screening measure is needed.

193. Rosenblu, S. et al. Davis-Eells Test Performance of Lower Class Retarded Children. Journal of Consulting Psychology, 19(1): 51-54, 1955.

This study concluded that for the subjects used the Davis-Eells Games did not reveal a "hidden intellectual potential not tapped by other intelligence tests." The possible reasons for this conclusion were that the Games were given in a culturally biased atmosphere and that the Games require abstract thinking ability which the "mentally deficient" subjects lacked.

194. Ruck, William W. Statistical, Legal, and Moral Problems in Following the EEOC Guidelines. 1972. 19p. ED 069 684 (MF-$0.65; HC-$3.29).

Statistical, legal and moral problems involved in following the EEOC guidelines are described. The guidelines require separate data for minority and non-minority groups with differential cut off scores for aptitude tests which have a racial bias. Problems reviewed include: identification of racial bias in tests is difficult; giving one race an advantageous cutoff over another may be unfair, creating legal challenges; and determining selection by race may diminish the effectiveness of the work group. The author suggests selection on the basis of proportion of numbers of each race applying, taking the top from each group.

195. Sampel, David D. The Academic Success of Black Students: A Dilemma. Journal of College Student Personnel, 12(4): 243-247, July 1971.

The efficiency of the two basic predictors (SCAT and high school rank) of academic success was investigated. Among the findings were that some of the predictors had little or no relevance for black males. A short review of research of studies dealing with the prediction of academic success for minority group students is included.

196. Saville, Muriel R. and Troike, Rudolph C. A Handbook of Bilingual Education. Washington, D. C., Center for Applied Linguistics, ERIC Clearinghouse for Linguistics, 1970. 69p. ED 035 877 (MF-$0.65; HC-$3.29).

The introduction to this handbook for teachers and administrators in bilingual education programs states: "Half of the children in the world are bilingual, and approximately one fourth of the people in the United States can communicate in more than one language. About ten percent of our population speaks a language other than English natively. For thousands of the children who enter school each year, English is a foreign language." The final chapter, stressing the importance of evaluation, discusses language and intelligence tests and home condition questionnaires.

197. Schools Council. "Cross'd with Adversity": the Education of Socially Disadvantaged Children in Secondary Schools. Working Paper 27. London, England, Evans Brothers Limited and Methuen Educational Limited, 1970.

Appendix 1 (pp. 131-133) discusses "The Place of Test Procedures and Some Notes on Educational Guidance" in relation to disadvantaged children. The point is made that standardized attainment tests can be used for the assessment of the effects of social deprivation and the planning of compensatory education but only in conjunction with other information. "To use tests without a clear recognition of their limitations is worse than not using them at all."

198. Schroeder, Glenn B. and Bemis, Katherine A. The Use of the Goodenough Draw-A-Man Test as a Predictor of Academic Achievement. Albuquerque, N. M., Southwestern Cooperative Educational Lab., 1969. 12p. ED 029 695 (MF-$0.65; HC-$3.29).

In an attempt to find a test which minimized cultural bias, three tests were administered to 335 first grade pupils. The subjects comprised 2 groups (123 Anglo children and 212 Spanish surnamed children). The Goodenough Draw-A-Man Test (GDAM) and the Lorge-Thorndike Intelligence Test (LT), Form A, were administered as measures of intelligence. The California Achievement Test (CAT), Form W, 1957 edition, was used as a measure of achievement. Results showed a mean IQ difference between the two groups of 4.39 on the LT and of 2.63 on the GDAM, with the Anglo group scoring higher on both tests. When the CAT scores were covaried with the LT IQ scores the only significant difference found between the two groups at the .01 level was in the "meaning of opposites" subtest. When the CAT scores were covaried with the GDAM IQ scores,

differences in six categories were revealed at the .01 level: (1) meaning of opposites, (2) total reading vocabulary, (3) total reading, (4) arithmetic problems, (5) total arithmetic reasoning, and (6) total battery. With this particular population, the GDAM and LT were nearly identical in predicting the CAT. The GDAM tends to bring these two divergent populations closer together as far as IQ scores are concerned. A bibliography and tables are attached.

199. Schultz, D. O. The Relationship Between College Grades and Aptitude Test Scores for Different Socioeconomic Groups. Research Bulletin 53-18. Princeton, N. J., Educational Testing Service, 1953.

It was concluded from the results of this study that SAT scores predict grades equally well and neither overpredict nor underpredict for all socioeconomic classes of college students.

200. Schwarz, P. A. Adapting tests to the Cultural Setting. Educational and Psychological Measurement, 23: 673-686, 1963.

This paper describes the methodology of test adaptation that evolved from earlier studies and from more recent findings in extending the work to other types of tests and to new populations.

201. Seiler, Joseph. Preparing the Disadvantaged for Tests. Manpower, 2(7): 24-26, July 1970.

Special instruction improved the test-taking skills of the culturally deprived and undereducated on employment tests and other commonly used screening devices.

202. Seiler, Joseph. Pretesting Orientation for the Disadvantaged: Experience in Manpower Experimental and Demonstration (E & D) Projects. 1970. 14p. ED 041 061 (MF-$0.65; HC-Not available from EDRS).

"Inasmuch as disadvantaged persons tend to have fairly restricted experience in taking employment and similar aptitude, intelligence and job knowledge

tests, it is reasoned that the lack of such exposure leaves them without test-making skills. ...pretest assistance [was] found useful in preparing disadvantaged persons to successfully compete on private employer and civil service job selection examinations." (p. 1)

203. Semler, J., and Iscoe, J. Comparative and Developmental Study of the Learning Abilities of Negro and White Children Under Four Conditions. Journal of Educational Psychology, 54: 38-44, 1963.

No overall race differences in learning ability was found between a Negro and a white group of children, the white group having shown significantly higher IQ's. This suggests that educators should exercise great caution in inferring learning ability from measured intellectual level alone.

204. Shaw, Ralph L. and Uhl, Norman P. Relationship Between Locus of Control Scores and Reading Achievement of Black and White Second Grade Children from Two Socioeconomic Levels. 1969. 11p. ED 036 575 (MF-$0.65; HC-$3.29).

This study investigates the effect of socioeconomic level (lower and upper-middle), race (black and white), and sex on locus of control of reinforcement scores, and the relationship between the latter scores and reading achievement in a sample of 211 second grade children. A stratified random sampling technique insured adequate levels of each dependent variable. Administered within a two-month period were: the paragraph meaning and word reading sections of the Primary I and Primary II levels of the 1964 Stanford Achievement Test; four subtests from Cattell's Culture Fair Intelligence Test; and the Bialer-Cromwell Children's Locus of Control Scale Analysis using a 2x2x2 factorial design found that, among the independent variables, only the main effect of socioeconomic level (SEL) and the interaction of SEL and race effects on locus of control scores were significant. The sole significant relationship between locus of control and reading scores was that white upper-middle SEL subjects having higher external scores had lower reading scores. Intelligence quotient was related to reading, but not to external scores in both SEL groups. The low reliability obtained using the Locus of Control scale indicates need for further development of this scale.

205. Shimberg, Benjamin. Prediction and Guidance. 1969. 8p. ED 044 425 (MF-$0.65; HC-$3.29).

Problems in the application and misapplication of test scores are discussed.

206. Shipman, Virginia C. Head Start Evaluation and Research Center, The University of Chicago, Report E, Comparative Use of Alternative Modes for Assessing Cognitive Development in Bilingual or Non-English Speaking Children. Chicago, Ill., University of Chicago, Head Start Evaluation and Research Center, 1967. 45p. ED 022 554 (MF-$0.65; HC-$3.29).

In order to assess the feasibility of alternative methods for determining the cognitive development of bilingual or non-English speaking children from a disparate cultural background, 28 Seminole Indian children from two Head Start centers were administered a series of intelligence tests, some based on verbal ability, some not. The tests used were (1) the Ravens Colored Matrices, (2) three Piagetian measures designed to assess the child's stage of concrete operation, (3) two measures of classificatory behavior (class inclusion and object sorting), and (4) the Stanford-Binet. The Ravens test purports to assess a person's present capacity for intellectual activity and has a minimum of verbal requirements. This test showed the least deficit for the Indian children and was a good predictor of the child's functioning on other tasks. The Stanford-Binet showed the children to be very deficient in intellectual ability. The children did poorly on the Piagetian measures, considerably below normative levels. On the sorting tasks, the Indian children scored below urban Negro children who had taken the same test. An appendix containing task descriptions follows this report.

207. Shuey, Audrey M. The Testing of Negro Intelligence, 2nd Edition. New York, N. Y., Social Science Press, 1966.

A review of research in narrative and tabular form is presented. 380 original investigations of Negro intelligence are reviewed. The research covers a period of approximately 50 years. Eighty psychometric tests, excluding different editions or forms of a given measure, each purporting to measure intelligence or some aspect of behavior related to it, were employed. Results include: 1) remarkable consistency in test results across all ages of "colored" subjects; 2) consistency across geographic lines in differences between Negroes and whites; 3) drop in average IQ of Negroes over past 20 years. The author concludes from these results the presence of native differences between Negroes and whites as determined by intelligence tests.

208. Skager, Roger A. Student Entry Skills and the Evaluation of Instructional Programs. A Case Study. Los Angeles, Calif., University of California, Center for the Study of Evaluation, 1969. 13p. ED 054 232 (MF-$0.65; HC-$3.29).

The findings of this study imply that teachers tended to direct instruction at skills which were relatively well-developed (as measured by an achievement test) at entry rather than at areas in which students were initially weak. Possible reasons why this might occur are examined and potentially important implications of these findings for instructional practice and evaluation methodology are discussed, especially with regard to instructional programs for educationally deprived students.

209. Solkoff, Norman. Race of Experimenter as a Variable in Research with Children. Buffalo, N. Y., State University of New York, 1971. 9p. ED 056 328 MF-$0.65; HC-$3.29).

The major aim of this study was to determine whether black and white children, ages 8-11, would differ in intellectual performance as a function of the race of the examiner. Two additional subgoals were: (1) to see the effect which the race of the examiner would have on the various subtest scores of the WISC; and (2) to find out if test anxiety scores would differ contingent on the race of the examiner. Four black and 4 white female high school graduates received extensive practical training in administering the WISC, and each then tested 14 black and 14 white children. The children all completed a test anxiety questionnaire prior to administration. Results show that, while the black children generally scored lower than whites, the predicted interaction between race of child and race of examiner was not obtained. In all instances, black examiners produced higher scores than whites. Finally, no significant main effects or interactions were found in relation to the anxiety measure. The findings are discussed.

210. Soares, Louise M. and Soares, Anthony T. Personality and Occupational-Interest Correlates of Disadvantaged Students. Paper presented at the Annual Meeting of the American Educational Research Assn., Chicago, Ill., April 1972. 7p. ED 061 285 (MF-$0.65; HC-$3.29).

Disadvantaged high school males indicated on the Minnesota Vocational Interest Inventory a preference for service-oriented jobs. Disadvantaged females tended toward office work and selling. On the Jr.-Sr. HSPQ form of the IPAT, both groups showed a pattern of personality which can be des-

cribed as group-dependent, restrained, relatively unfrustrated, self-assured, feeling few obligations, somewhat detached yet self-assured, less oriented to intellectual pursuits, and a tendency toward emotional instability and low-ego strength. The disadvantaged females were significantly differentiated from the males in tough-mindedness and self-reliance. The correlation of these measures was .58, significant at the .01 level.

211. Stanley, J. C. and Porter, A. C. Correlation of Scholastic Aptitude Test Score with College Grades for Negroes Versus Whites. Journal of Educational Measurement, 8(4): 199-217, 1967.

This study found that the SAT correlated well (+.25 to +.45) with freshman grades in black colleges. This evaluation was based on a comparison with SAT correlations in predominantly white colleges in the same state. The conclusion was reached that SAT-type tests were valid as predictors for blacks who were competing with blacks taught by black instructors.

212. Stanley, Julian C. Predicting College Success of Educationally Disadvantaged Students. Report No. 79. Baltimore, Md., The Johns Hopkins University, The Center of the Study of Social Organization of Schools, 1970. 45p. ED 043 295 (MF-$0.65; HC-$3.29).

From a review of the literature the author concludes that admission to selective colleges should be based substantially on test scores and high-school grades, whether or not the applicant is from a minority racial, ethnic, or socioeconomic group. The gap between the academic promise of educationally disadvantaged applicants and the usual minimum demands of the institution should not be greater than explicit provisions for remediation, tutoring, coaching, and perhaps curricular reform can bridge.

213. Stodolsky, S. S. and Lesser, G. Learning Patterns in the Disadvantaged. Harvard Educational Review, 37(4): 546-593, 1967.

"The types of achievement and intelligence tests which are most often used can only have limited value in describing the cognitive functioning of children. In almost all instances we are concerned with scratchings on an answer sheet not with the ways in which a student arrived at a conclusion. No matter how much we may think we know by looking at scores on such psychometric procedures, unless they are constructed to reveal reasoning processes we simply won't get at them." (p. 15)

214. Stone, Chuck. Psychology and the Black Community: From Arthur, 1853 to Arthur, 1969. 1971. 23p. ED 054 286 (MF-$0.65; HC-$3.29).

"The use of tests as the principal and frequently only assessment instrument to determine intellectual proficiency, educational accessibility, job placement and professional advancement is a widespread racist cancer in the American body politic." (p. 18) "...tests must not be permitted to help maintain educational concentration camps or culture ovens for Black and Spanish-speaking children." (p. 22)

215. Stone, Chuck. Testing and the Educational Power Struggle. Integrated Education, 9(4): 4-10, July-August 1971.

This article discusses the testing experience as it affects the black community. It contends that nationally standardized tests have: 1) excluded the black population in establishing their norms; 2) become instruments for reinforcement of negative black images; and 3) been used inadvertently or otherwise to bar blacks from higher education.

216. Taba, Hilda and Elkins, Deborah. Teaching Strategies for the Culturally Disadvantaged. Chicago, Ill., Rand McNally and Co., 1966.

"A basic cause of defective teaching strategies, of ineffective selection of content and materials, of inadequate approaches, and of poor learning atmosphere is lack of systematic, all-encompassing, and continuing diagnosis... Teaching in the dark is a questionable practice in any case. This is especially true in teaching students whose background, motivation, learning styles, and habits deviate from the usual, and whose attitudes toward school and teachers have been distorted by years of frustration stemming from unsuccessful learning careers. It is important that this diagnosis encompass a wide array of types of learning: knowledge, cognitive styles, levels of cognitive development, emotional dynamics, self-concept, skills, attitudes, and interests, among others." (p. 23)

217. Temp, George. Test Bias: Validity of the SAT for Blacks and Whites in Thirteen Integrated Institutions. Princeton, N.J., Educational Testing Service, 1971. 18p. ED 052 248 (MF-$0.65; HC-$3.29).

Differential prediction of grade point average for black and white freshman

students was empirically investigated at 13 integrated institutions by comparison of regression planes. Particular attention was given to the possibility that prediction procedures that are appropriate for white (majority) students would under-predict the performance of black (minority) students. The data tend to support, among others, the following generalizations: (1) a single regression plane cannot be used to predict freshman GPA for both blacks and whites in many of the institutions studied; (2) nevertheless, if prediction of GPA from SAT scores is based upon prediction equations suitable for majority students, then black students, as a group, are predicted to do about as well as (or better than) they actually do. Analysis demonstrated that a general conclusion applicable to all institutions is not justified. Admissions officers are urged to consider and conduct institutional self-studies routinely on the question of differential predictive validity.

218. Thomas, Alexander. Retardation in Intellectual Development of Lower-Class Puerto Rican Children in New York City. Final Report. New York, N.Y., New York University Medical Center, 1969. 64p. ED 034 370 (MF-$0.65; HC-$3.29).

Examiner testing style did appear to affect total score for bilingual children.

219. Thomas, C. L. and Stanley, J. C. Effectiveness of High School Grades for Predicting College Grades of Black Students: A Review and Discussion. Journal of Educational Measurement, 6(4): 203-215, 1969.

The authors reexamine and analyze data from previous studies and find that high school grades do not consistently make the greatest contribution in the prediction of college grades for black students. They conclude by stating that academic aptitude and achievement tests, in comparision to high school grades, are better predictors of college grades for blacks than they are for whites. They also feel that the best results can be obtained when both test scores and high school grades are used to predict college grades.

220. Thorndike, Robert L. Concepts of Culture-Fairness. Journal of Educational Measurement, 8(2), 63-70, Summer 1971.

Fiarness of a test relates to fair use. It is shown that when 2 groups differ appreciably in mean test score, the test is "unfair" to the lower group as a whole in the sense that the proportion qualified on the test will be smaller, relative to the higher-scoring group, than the proportion that will reach any specified level of criterion performance.

221. Torrance, E. Paul. Are the Torrance Tests of Creative Thinking Biased Against or in Favor of "Disadvantaged" Groups? Gifted Child Quarterly, 15(2): 75-80, Summer 1971.

This paper discusses various studies which used the Torrance Test of Creative Thinking over a period of years. Little difference was found between races or between socioeconomic groups. In some cases it was the "disadvantaged" which scored higher. However, creativity must be given a chance to develop. To provide this channel, school and community programs have to respect the creative positive of disadvantaged children and build upon them rather than stress compensation for deficits.

222. Trider, Mary S. The Right to Read and Standardized Testing: A Necessary Dimension, Reading Teacher, 24(4): 320-330, 368, January 1971.

This article asserts that standardized tests used in Title I, ESEA programs can provide valuable information concerning the children and the quality of instructional services in a particular area: trends and tendencies within school instructional programs, strengths and weaknesses of various approaches involved, individual needs and difficulties of student populations, specific areas that need to be further investigated, inconsistencies between ability and achievent scores that merit more study. (pp. 320-321)

223. Tumin, Melvin M., ed. Race and Intelligence. New York, N. Y., Anti-Defamation League of B'nai B'rith, 1963. 56p. ED 001 938 (MF-$0.65; HC-$3.29).

The report of a UNESCO committee stated that there were not measurable differences between what is due to innate capacity and what is the result of environmental influences, training, and education. Tests which have made allowance for most differences in environmental opportunities have shown essential similarity in mental characteristics among all human groups. The UNESCO report has been challenged by a series of publications contending that it ignored scientific evidence that the Negro in the United States is inferior in intelligence to the white. Four distinguished experts in tests and measurements, psychology, social and cultural patterns, and anthropology were asked to state their opinions on the challenge. Questions were asked on innate capacity, test reliability, Negro Intelligence, and other innate differences between Negro and whites. The scientists are in substantial agreement that the claims of the challenging publications cannot be supported by any substantial scientific evidence. Morever, future claims of

innate differences between Negro and white cannot be valid unless: distinctive racial homogeneity of the group is demonstrated, social and cultural backgrounds of the group are equal, and adequate tests of proved reliability and validity are used.

224. Turnbull, William W. Meeting the Measurement Needs of Education. 1971. 17p. ED 049 309 (MF-$0.65; HC-$3.29).

An interlocking, coordinated program involving guidance, testing, admissions, financial aid, curriculum and research is proposed as a comprehensive and relevant measurement requirement for the children of poor and minority group backgrounds.

225. Turnbull, William W. Socio-economic Status and Predictive Test Scores. Research Memorandum 51-9. Princeton, N. J., Educational Testing Service, 1951.

The author concludes by saying that "fairness, like its amoral brother, validity, resides not in tests or test scores but in the relation of test scores to criteria."

226. Turner, Charles. Effects of Race of Tester and Need for Approval on Children's Learning. Journal of Educational Psychology, 62(3): 240-244, June 1971.

Eighty white male children were given a simple serial learning task by experimenters of different racial identities. The experimenters motivated the children with verbal approval during the testing. It was found that those subjects tested by same race male testers performed significantly better than subjects tested by different race male experimenters.

227. Underwood, M. Hyre. Environmental Factors That Influence Achievement in a Reading Improvement Program. Whitewater, Wisc., Wisconsin State University, 1969. 56p. ED 052 920 (MF-$0.65; HC-$3.29).

Including tests which measure environmental factors in a diagnostic reading test battery, teachers should more adequately meet the needs of the children in Title I programs and be able to approach reading difficulties on a broader scope.

228. Vernon, P. E. Intelligence and Cultural Environment. London, England, Methuen, 1969.

Part IV of this book discusses a battery of tests used in a variety of studies conducted by the author. The chapters are: "The Tests," "What the Test Measure," "The Influence of Background and Personality Factors on the Scores of English Boys," Testing in the Hebrides." Part V discusses cross-cultural studies. Among the areas covered is the use of tests in the studies.

229. Walker, Wanda. Development of a Readiness Test for Disadvantaged Pre-School Children in the United States. Final Report. Maryville, Mo., Northwest Missouri State College, 1969. 147p. ED 037 253 (MF-$0.65; HC-$6.58).

The object of this project was to develop a culture-fair, non-verbal individual readiness test for disadvantaged preschool children. Two equivalent forms of the test were developed so that teachers could administer the first form to identify needs early in a preschool program, set up specific remedial programs for individuals, and use the second form of the test to assess the efficiency of the program and individual progress. A total of 6662 children in 364 Head Start and Day Care Centers were tested to refine and standardize Form A of the Test. Analysis and norming of Form B was carried on in 301 centers with 5271 children participating. Item analyses of both forms of the test were made. The data indicate that both forms of the test are sufficiently valid and reliable for use in an evaluation program for disadvantaged preschool children.

230. Walker, Wanda. Walker Readiness Test for Disadvantaged Pre-School Children; Forms A and B. Washington, D.C., Office of Education (DHEW), Bureau of Research, 1969. 28p. ED 045 736 (MF-$0.65; HC-$3.29).

This instrument assesses readiness for school in preschool disadvantaged children. Two forms, A and B, are available and it is recommended that one be used as a diagnostic pretest and other as a progress measure. The test is an individually administered, non-verbal, untimed, but paced instrument. The child responds to questions, given in English, French, or Spanish, as appropriate, by pointing to one of four pictures or figures. The instrument tests the child's ability to recognize similarities, differences, the cardinal number of a set, and missing parts. No special equipment or training is necessary to administer. The teacher records responses on a separate answer sheet which is hand scored using a stencil. A copy of the answer sheet and stencil are included. The mean, median, standard deviation, quartile ranks, and percent-

ile ranks for the norming groups are given in 6 month intervals from age 4.0 to 6.5. In validity studies, the average correlation of Form A with 17 commonly-used, standardized readiness and IQ tests was found to be .48; Form B, .53. The correlation with Head Start and day care teachers' ranking of readiness for entering public school was .68 for Form A and .63 for Form B. Split-half reliability on Form A was .87 and on Form B, .89. Parallel form reliability correlations for several populations are available.

231. Wardrop, James L. Towards a Broader Concept of Educational Assessment. Urbana, Ill., University of Illinois, Center for Instructional Research and Curriculum Evaluation, 1972. 28p. ED 060 062 (MF-$0.65; HC-$3.29).

This paper illustrates how changing emphases and trends in education have led to a reformulation of notions of assessment and an awareness of the limitations of current measurement technology to deal with newer conceptualizations. Discussed are: 1) the antecedents of learning and the transactions which take place during the learning process, as well as with outcome measures; 2) the need to look at intentions; and 3) the concern with a plurality of goals and values as these relate to judgments based on assessment data. Problems highlighted included criterion-referenced measurement, the measurement of change, and the limitations of standardized tests as instruments for accountability.

232. Wasik, Barbara H. and Wasik, John L. Performance of Culturally Deprived Children on the Concept Assessment Kit--Conservation. Child Development, 42(5): 1586-1590, November 1971.

The Concept Assessment Kit--Conservation was administered to 117 children in an ungraded primary school for culturally deprived children. The scores were compared with the norm population and the results indicated that the experimental group was taking from 1 to 2 years longer to attain the level of acquisition of conservation demonstrated by the norm group. It is suggested, on the basis of the data obtained in this study, that those who work with culturally deprived populations recognize discrepancies in cognitive development between children of different cultural backgrounds and, being able to test for them with a fairly quick measurement, be able to better plan realistic learning experiences for the children.

233. Weiner, M. and Feldmann, S. Validation Studies of a Reading Prognosis Test for Children of Lower and Middle Socio-Economic Status. Educational and Psychological Measurement, 23(4): 807-814, 1963.

The purpose of these studies was to determine whether a reading prognosis test could be constructed to measure future reading ability based on present skills and knowledge of children from different socio-economic levels. It was concluded that poor readers from any socio-economic status group can be identified before formal training in reading takes place and that their skill deficiencies underlying reading can be ascertained.

234. West, Lloyd Wilbert. Assessing Intellectual Ability with a Minimum of Cultural Bias for Two Samples of Metis and Indian Children. Edmonton, Alberta, University of Alberta, 1962. 139p. (Available from Western Microfilm Limited, Edmonton, Alberta, Canada).

An investigation was designed to ascertain the effects of cultural background on selected intelligence tests and to identify instruments which validly measure intellectual ability with a minimum of cultural bias. A battery of tests, selected for factor analytic study, was administered and replicated at four grade levels to a sample of Metis and Indian children in Canada. The tests employed generally included conventional intelligence tests and subtests as well as several promising non-verbal or "culture-reduced" varieties. Some tests showed significantly less cultural bias than others. The "culture-reduced" tests identified showed significantly less increase with grade level as a result of selection and school treatment than traditionally used tests as well as substantial correlations with academic achievement. It was found that the degree of cultural bias may vary in a test at different levels. A longitudinal study, intensive item analysis of tests, and a recombination of sub-tests or items were recommended.

235. Whiteley, John M. and Wientge, King M. Workshop for Counselors and Educators Concerned with the Education, Training and Employment of Minority Youth. Final Report, Part II, Discussion Guide to the Problems of the Culturally Deprived: An Introduction for Teachers and Counselors. University College Research Publications Number 8. St. Louis, Mo., Metropolitan St. Louis Plans for Progress Companies and Washington University, 1965. 80p. ED 021 957 (MF-$0.65; HC-$3.29).

This report contends that intelligence test scores underestimate intelligence of culturally disadvantaged children. Discussed are some general considera-

tions of the testing of minority group youth. Various limitations of standardized tests for psychological assessment are stressed.

236. Whitesides, Sheila. Administering the POSR to Lower Class Children. 5p. 1965. ED 011 608 (MF-$0.65; HC-$3.29).

The children tested were of Latin-American descent and from a low socioeconomic status. The investigator concluded: 1) reading questions in both English and Spanish may help at all grade levels; 2) Spanish-speaking persons can elicit more cooperation from the pupils, and, perhaps, more honesty in answers than an Anglo.

237. Williams, E. Belvin. Testing of the Disadvantaged: New Opportunities. 1971. 15p. ED 054 282 (MF-$0.65; HC-$3.29).

This essay discusses testing of the disadvantaged and the possibilities for new approaches to the problem. The black community's resistance to testing and the legal implications of employment and educational discrimination which results from testing are discussed. The objections to testing the disadvantaged are reviewed. At the end, it is hoped that other criteria than test scores will be used for university admission and employment.

238. Williams, R. L. Black Pride, Academic Relevance and Individual Achievement. Counseling Psychologist, 2(1): 18-22, 1970.

"Measured IQ in Black people is not an accurate measure of intelligence... [because]...measures of individual achievement and the conditions confronting the learner are different for Blacks and Whites." (p. 19) The author discusses the bias of conventional tests and contends that these must be "revalidated to include Black responses to white-oriented tests." (p. 20).

239. Yater, Allan C. et al. Goodenough-Harris Drawing Test and WPPSI Performance of Disadvantaged Preschool Children. Perceptual and Motor Skills, 33(3): 967-970, December 1971.

The purpose of this study was to investigate and compare the performance of a group of disadvantaged preschool children in the Goodenough-Harris Drawing

Test and the Wechsler Preschool and Primary Scale of Intelligence. Analysis of the performance of 48 disadvantaged preschool children on the two tests showed that both tests yield IQ estimates below the respective norm groups for both instruments. Goodenough-Harris IQs were in the dull-normal range, while WPPSI IQs were in the normal range.

240. Yourman, Julius. The Case Against Group I.Q. Testing in Schools with Culturally Disadvantaged Pupils. In: The Disadvantaged Child, Issues and Innovations, edited by Joe L. Frost and Glenn R. Hawkes. Boston, Mass., Houghton Mifflin, 1966.

The decision of the New York City Board of Education to discontinue the use of group intelligence testing is discussed. A brief history of intelligence testing is presented. It is suggested that achievement measurements be substituted for intelligence tests and an effort be made to individually assess each child's potential.

241. Zach, Lillian. Current Thought on Intelligence Tests. Psychology in the Schools, 3(2): 116-123, April 1966.

"Unfortunately, in many cases, the predictive validity has been 'built in' by our very faith in it. That is, the IQ will frequently determine the kind of educational environment which will, in turn, support the IQ. This has often been the case in dealing with socially disadvantaged children (Clark 1963)...prediction of true potential can have little meaningful validity unless the previous and subsequent life experiences of the individual are studied." (p. 118)

242. Zach, Lillian. The IQ Debate. Today's Education, 61(6): 40-43, 65-66, 68-69, September 1972.

This article gives a brief history of intelligence testing and enters a plea for a better use of testing to meet the individual needs of students.

243. Zirkel, Perry Alan and Greene, John F. The Academic Achievement of Spanish-Speaking First Graders in Connecticut. Hartford, Conn., Connecticut State Department of Education, 1971. 8p. ED 054 275 (MF-$0.65; HC-$3.29).

This study sought to demonstrate that deficiencies Puerto Rican children show in verbal ability and academic achievement might not exist if initial instruction and testing were in Spanish. Puerto Rican first-graders scored significantly lower on English forms than all other ethnic groups and significantly higher on the Spanish forms.

Subject Index

(By citation number)

Ability Grouping

28, 43, 67, 68, 69, 145, 154

Ability Identification

18, 19, 25, 26, 43, 77, 82, 83, 91, 93, 96, 110, 122, 129, 143, 145, 152, 180, 201, 203, 233

Academic Achievement

22, 27, 28, 58, 74, 104, 119, 123, 128, 164, 195, 198, 204, 238, 243

Achievement Tests

40, 63, 87, 240

Adults

162

American Indians

24, 143, 151, 153, 206, 234

Aptitude Tests

11, 31, 33, 42, 83, 85, 147, 148, 161, 199, 210

Bibliographies

64, 111, 152, 167, 172, 189

Bilingualism

9, 12, 21, 167, 189, 196, 206, 218, 236, 243

College Placement

15, 25, 44, 56, 90, 124, 131, 237

Criterion Referenced Tests

81, 111, 117, 169, 181, 231

Cultural Differences

8, 10, 14, 39, 55, 57, 73, 101, 103, 130, 134, 137, 200, 223, 232

Cultural Factors

15, 17, 31, 32, 36, 37, 46, 70, 74, 79, 80, 85, 89, 119, 120, 121, 122, 139, 143, 149, 155, 180, 185, 214, 216, 220, 228, 234

Culture Free and Culture Fair Tests

4, 5, 9, 10, 13, 17, 33, 34, 36, 39, 50, 51, 60, 101, 105, 130, 134, 142, 150, 173, 190, 193, 229, 234

Davis-Eells Games

4, 193

Discrimination

2, 5, 54, 59, 64, 67, 71, 75, 82, 89, 94, 95, 98, 100, 138, 148, 154, 188, 214, 215, 220, 237

Educational Accountability

45, 59, 74, 111, 224, 231

Educational Diagnosis

24, 91, 102, 106, 117, 137, 171, 187, 197, 205, 208, 216, 222, 227, 230, 232 240

Educational Objectives

18, 38, 43, 45, 59, 83, 90, 208, 216, 224

Goodenough Draw-A-Man Test

198, 239

Higher Education

41, 42, 44, 53, 56, 75, 77, 90, 102, 107, 108, 164, 177, 184, 195, 199, 211, 212, 217, 219

Intelligence Tests

3, 4, 14, 16, 17, 26, 30, 36, 37, 46, 52, 57, 61, 62, 76, 84, 101, 112, 113, 116, 118, 119, 120, 122, 125, 135, 144, 158, 165, 179, 191, 196, 206, 207, 213, 235, 239, 240, 241, 242

Intervention

32, 57, 87, 105, 123, 168, 173, 186, 197, 212, 222, 227

Johns Hopkins Perceptual Test

190, 191, 192

Language Ability

40, 78, 89, 115, 133, 153, 178, 243

Language Development

34

Language Handicaps

6, 9, 21, 24, 35, 87, 135, 180, 183, 187

Language Tests

34, 78, 167, 196

Mental Retardation

35, 135, 154, 155, 193

Metropolitan Tests

12, 114, 115, 126, 159, 160

Mexican Americans

35, 58, 73, 99, 109, 114, 149, 154, 155, 159, 189, 236

Nonverbal Tests

80, 147, 229, 230, 234

Objective Tests

110

Peabody Picture Vocabulary Test (PPVT)

1, 20, 48, 49, 125, 158, 191, 192

Performance Tests

23

Prediction

15, 22, 26, 32, 41, 42, 44, 56, 58, 65, 70, 75, 84, 94, 106, 107, 108, 113, 124, 129, 131, 136, 141, 143, 164, 177, 195, 198, 199, 205, 211, 212, 219, 225, 233, 241

Predictive Validity

159, 160

Preschool

6, 48, 49, 106, 116, 125, 168, 170, 171, 183, 186, 190, 191, 192, 206, 229, 230, 239

Psychological Tests

7, 8, 79, 100, 140, 157, 235

Puerto Ricans

6, 9, 182, 189, 218, 243

Racial Differences

3, 11, 40, 47, 55, 65, 66, 91, 119, 120, 127, 142, 165, 178, 203, 207, 209, 219, 221, 223, 238

Racial Factors

1, 29, 30, 53, 86, 102, 109, 118, 128, 132, 166, 176, 177, 184, 194, 204, 214, 226

Self Concept

19, 71, 86, 95, 100, 115, 145, 146, 153, 210

Social Factors

3, 67, 82, 83, 93, 94, 95, 104, 107, 140, 146, 155

Socioeconomic Influences

37, 51, 52, 53, 61, 74, 86, 104, 109, 119, 121, 126, 129, 130, 139, 142, 144, 148, 149, 150, 151, 156, 161, 163, 165, 166, 185, 204, 221, 225, 233

Standardized Tests

20, 25, 28, 52, 53, 70, 72, 77, 84, 90, 92, 93, 96, 97, 98, 103, 115, 123, 133, 136, 153, 162, 163, 164, 168, 170, 176, 185, 187, 197, 207, 215, 222, 231, 235

Stanford Binet

48, 125, 158, 174, 178, 183

Test Bias

11, 13, 14, 22, 41, 42, 54, 60, 61, 62, 63, 64, 69, 71, 72, 78, 85, 98, 99, 109, 121, 126, 133, 134, 136, 137, 138, 141, 144, 149, 155, 156, 163, 185, 194, 200, 214, 217, 234, 238

Test Construction

18, 23, 50, 51, 61, 72, 79, 80, 99, 112, 117, 139, 144, 179

Test Criticism

16, 71, 92, 94, 96, 138, 140, 142, 157, 172, 213, 214, 215, 237, 242

Test Examiners

1, 29, 30, 47, 48, 66, 88, 97, 118, 127, 166, 176, 178, 184, 209, 218, 226, 236

Test Interpretation

2, 5, 7, 10, 15, 16, 19, 24, 37, 45, 63, 76, 81, 113, 136, 146, 151, 154, 156, 163, 176, 179, 181, 182, 203, 213, 225

Test Reliability

12, 14, 68, 112, 114, 151, 159, 181

Test Results (Use of)

38, 62, 76, 92, 96, 97, 103, 124, 137, 150, 156, 157, 169, 186, 205, 208, 213, 215, 220, 223, 228, 231, 242

Test Selection

23, 38, 50, 152

Test Validity

2, 8, 12, 13, 21, 22, 31, 49, 54, 68, 70, 73, 75, 81, 112, 113, 114, 118, 131, 139, 151, 156, 174, 176, 181, 188, 192, 211, 217, 225, 241, 242

Test Wiseness

48, 55, 116, 170, 201, 202

Testing Conditions

29, 75, 88, 116, 127, 176, 193, 226

Torrance Tests of Creative Thinking

27, 221

Vocational and Employment Testing

3, 31, 64, 65, 132, 147, 188, 194, 201, 202, 210, 237

Wechsler Adult Intelligence Scale (WAIS)

20

Wechsler Intelligence Scale for Children (WISC)

46, 88, 166, 209

Wechsler Pre-School Primary Scale of Intelligence (WPPSI)

174, 239